

Another monopoly on the block

This month's sale of the telephone company British Telecom will not bring competition but may net the government £4,000 million. The objectives are admirable but the methods are wrong.

JUST as there is no general answer to the question how long is a piece of string, so there is no easy way of telling what a national telephone network may be worth. That may be the most charitable excuse for the stratagems to which the British Government has gone, in the past few months, for selling off 51 per cent of the shares in its nationalized telecommunications industry, now trendily called British Telecom. Last week, after an advertising campaign intended to interest men in the street in becoming shareholders (and which must have done more for general business education than decades of formal education), British Telecom and its merchant bank last week published the draft of the prospectus on which its shares will be sold on 28 November. The only missing ingredients in the prospectus are the price at which shares will be offered for sale, and the precise proportions that will be set aside for employees of the company, investors overseas and so on. Everybody is a little agog that on this one day, anything up to £4,000 million may change hands (although only 40 per cent of it will be payable on the nail).

While there is every reason to applaud the government's chief objective, that of introducing some of the self-balancing discipline of the marketplace into the regulation of a gigantic industry, suspicion must persist about its secondary objectives (balancing its budget by selling off its capital assets) and there are grave doubts about its methods.

The laudable primary objective is now quite widely shared. Only this year, the US telephone network, AT&T, has been broken up into a number of smaller units, with geographically restricted monopoly rights. The Japanese Government is dickering with similar schemes. Everywhere, monopolistic telecommunications have become technically backward (or less technically advanced than they should be) and commercially slothful. Two quite separate tendencies have made the telephone monopolies indefensible. First, as experience in the United States has shown, allowing small companies to offer "value added services" by renting trunk circuits from the monopoly and selling on their use can benefit both the entrepreneurs and their customers. (The monopolies protest that these parasitic middlemen are merely skimming the cream off their legitimate business, to which the answer is that they should design their tariffs more intelligently.) Second, technology has often left the monopolies standing almost still and, in the past decade, with satellite communications in particular, has begun to raise the question whether telephone systems need now be the monopolies that seemed inevitable in the days of Alexander Bell. The British Government deserves some praise for recognizing that the time has come for change.

The British solution to the problem of the monopoly is at best half-hearted. There are three innovations. First, half of the telecommunications will in future be owned by private investors, not the government. The chief practical consequence will be that British Telecom will in future be able to raise capital from private sources without embarrassing the Treasury by inflating what is called the Public Sector Borrowing Requirement (which will assist technical innovation). Especially because the government promised, in last week's prospectus, not to vote its shares in ordinary circumstances, this change by itself is not a safeguard for the users of the network but, if anything, a device whereby the new owners can decide what profit they should be making. To guard

against exploitation, there are two safeguards — a regulatory office (called Oftel) with powers to rule on complaints of unfair competition which has yet to show its teeth (but they look sharp) and a requirement imposed by the government that domestic telephone tariffs should not in the next five years increase by more than two per cent less than the inflation rate (which is meaningless, because the new corporation is almost certain to wish to finance a greater share of its development by borrowing than in the past). Third, there is a simulacrum of independent competition, a quite separate telecommunications network called Mercury.

The best that can be hoped for, in these novel circumstances, is that the competition will prosper. The worst possible outcome is unthinkable, for the political trouble there would be if Mercury should be in danger of collapse would have the government and the Bank of England at its side in a flash. (Perhaps the moral of last week's prospectus is to invest in Mercury, not British Telecom.) Meanwhile, the British Telecom prospectus cheekily reminds potential investors that the restriction of its tariffs laid down by the government will apply to only 45 per cent of its business. (To be fair, the rest includes international telephone traffic, where competition is likely in due course to be fierce, and the sale of equipment, where it is already cutthroat.)

Meanwhile, the question remains how much the business may be worth. The government's embarrassment is that there is no objective way of putting a value to its worth. The value of the income stream from a monopoly enterprise is not necessarily related to anything but the decisions taken by individuals as to what the profit should be. The value of the network's assets, calculated from the cost of providing them in the first place but abated by depreciation, may conceal as assets equipment so inefficient that, in a competitive world, it would be thrown away. One of the curious anomalies in last week's prospectus is the brief statement on depreciation policy, disclosing that British Telecom reckons on an average life of between 6 and 20 years for its old Strowger electromechanical equipment, but a uniform ten years for new digital equipment.

The sale now arranged has inevitably an air of unreality. Ultimately, what will matter is how much people are prepared to pay for half of British Telecom. The best way to have found out would have been to put the shares on sale in dribs and drabs, on exactly the pattern in which the government itself sells government stock. The way in which things have been arranged, however, is certain to cause embarrassment. If the shares on offer are not taken up, the government will be laughed at. If on the other hand, the shares sell like hot cakes, it will be accused of having given away a public asset. It has only itself to blame. □

Research competition

British universities had better become knowledgeable about their research.

THOSE who think that the long agony of the British academic research enterprise must be coming to an end had better think again. That is the simple lesson to be drawn from last week's announcement by the Science and Engineering Research Council that it is now embarked on finding fields of research or important



research institutions, from which it can withdraw support. The plan is to find things to drop by early next year, and no later than April. Meanwhile researchers who may with gratitude have been allowed to retire from academic committees brooding on the future of their universities will now find themselves sucked up into another round of endless meetings, this time about research support. Before the dust has settled, there will be a great deal to be said on the important and daunting questions underlying the new round of enquiries.

In the past few years, the much-vaunted British system for supporting academic institutions, the dual-support system, has come to resemble a pair of millstones between which the research enterprise is being ground. Theoretically of course, there should now be no immediate problem. The contraction of the universities has come to an end, the British Government has been hinting that henceforth "level funding" will apply, while the government has in strict monetary terms kept its prime minister's promise (in 1980) that the civil research budget would be "protected". The snag is that constant budgets for the universities in present circumstances imply a steady drain of 0.5 per cent a year in real resources, while the constant research budget (spent through the research councils) has now to cover a greater range of activities, including payment of pensions to people whose services are no longer required. No wonder that the University Grants Committee is threatening (but some will say promising) greater selectivity in support of particular universities, and that the research councils are having to draw in their horns.

With this in prospect, any sensible British university should take a few simple protective steps. First, it should set up a research committee to attempt to decide on its priorities. The inescapable reason for doing this is simply that the universities were told to do just this two years ago, in the report of an investigation in whose existence the grants committee had a hand. But there is also a chance that such an arrangement would give the universities a capacity which they at present lack for knowing and then for understanding the range of research that their academics at present conduct. (In passing, it might be acknowledged that part of the trouble underlying British academic research in recent decades has been the indifference of universities to their own achievements.) And there is no doubt that such an understanding will be essential eighteen months from now, when the men from the University Grants Committee set out with their scalpels.

To be safe, however, universities should be careful to distinguish between academic research and the now fast-growing volume of industrial contract research carried out within universities (the once watchful eyes of the revenue authorities temporarily blinkered by the climate of the times). To say this is not to echo old-fashioned notions about the relative merits, intellectual and otherwise, of the two kinds of activities but is simply to say that managing academic and industrial research are different activities, often calling for different people. And while there is every sign that the British Government (and thus, to some extent, the grants committee) will applaud those institutions that earn a large proportion of their income from contracts, there is as yet no sign that the grants committee will relax its traditional standards of scholarship when choosing between the sheep and the goats. Moreover, that process cannot now be long delayed.

For the institutions responsible for managing the impending change, the grants committee and the research councils, the benefits of acting quickly are now overwhelming. The grants committee wants to concentrate research at a smaller number of universities within a system accommodating as many students as at present, while the research councils must somehow shed a large proportion of their fixed overhead cost, that represented by their commitment to research institutes and the people there employed. In an ideal world, they would put their heads together and recognize that their problems would more easily be solved in concert than separately. The only evil consequence would be that the goats (or is it the sheep?) among the universities would more quickly know where they stand — and that a greater proportion of them would be wrongly classified. Such a recognition of common cause may be the next threat to the universities. □

Animal crackers

Using the cloak of animal rightists' groups to persecute people is wrong.

IN the pathology of personal vindictiveness, the way in which an unidentified resident of Pennsylvania has apparently set out to damage Dr Robert Weinberg, a professor at the Whitehead Institute at Massachusetts Institute of Technology, breaks alarming new ground. Last month, researchers in Britain as well as the United States were circulated with an anonymous notice, purportedly from an organization called the Animal Liberation Front, which uttered unspecified threats at those in cancer research who use animals in their experiments, but which singled out Dr Weinberg for having presented at *Nature's* September conference in Boston a paper supposedly encouraging painful animal experiments. The title of the paper was given as "Neoplasms as uncontrolled regenerative hyperplasia", which will have seemed a little odd to the many who know Dr Weinberg's work. To suggest that his activities entail more than the most marginal dependence on laboratory animals is, of course, laughable, which is the spirit in which the incident was reported a few weeks ago (*Nature* 4 October, p.401).

Since then, the plot has thickened. Whoever circulated the anonymous notice seems also to have been circulating scientific papers supposedly written by Weinberg. One curious feature of these documents is that they appear not to involve the kinds of experiments with animals that would send the animal welfare people into a frenzy. Weinberg has discovered that the fake paper he was said to have read at the *Nature* conference (at which, in reality, he gave a sober review of the recent molecular biology of oncogenes) had actually been mailed to him, under different authorship, several months earlier (when it had been filed together with other crank cancer theories). More mystifying is a letter, apparently widely circulated, that purports to be a manuscript submitted to the *New England Journal of Medicine* which puts forward a proposal for the use of steroid hormones in the treatment of breast cancer. An accompanying note says that the editors of the *Journal* would be grateful for comment. Again, there is no mention of the use of experimental animals.

There are several lessons to be learned from this curious chain of events, not the least of which is the discovery of how easily a researcher's attention may be distracted by malevolent and apparently pointless activity. It is, however, intolerable that productive people should be thus preoccupied. The best way to deal with nuisance mail of this kind is to tell the police, urge them to investigate and to forget about it all. That, unfortunately, is easier said than done, as will be verified by anyone who has personal experience of knowing there is somebody nursing some kind of hate. The calculation that the anonymous watcher is probably psychiatrically disabled is little comfort. Since Weinberg has not been the only target of the anonymous nuisance-monger's attention, it may be worthwhile that others affected should communicate with him or with *Nature*.

Second, while there is no serious danger that a well-established reputation could be shaken by the circulation of nonsensical papers, there is every chance that people's professional opinions could be altered in subtle ways by the circulation of, say, preprints masquerading as serious works. People should be on their guard against such dangers, unlikely though they may appear.

The use of the slogan "Animal Liberation Front" on some of the communications from Pennsylvania points to another hazard — the danger that ill-considered actions by people who should know better will provide a cloak beneath which unwelcome allies may have their fling. Of necessity, because it is a banner for often illegal activities, the Animal Liberation Front cannot be a formal organization, with a constitution and a membership list. The result is that anybody can join, gaining attention for private causes from the publicity that has been engendered by those whose interest is to break into laboratories and animal houses. The hazard is one of those attending the activities of all terrorist organizations, however well-intentioned. □

US space research

Space station threat to science spending

AN enormous share of the US National Aeronautics and Space Administration (NASA)'s space science budget may wind up supporting the space station, according to some initial assessments by scientists advising the agency. Peter Banks of Stanford University, chairman of a panel looking into scientific uses of the station, said last week that \$400–600 million a year might be spent by "code E" — NASA's speak for its science budget — on space station projects. The NASA science budget is now \$1,404 million, which goes to support both basic research, mainly at universities, and such major projects as the space telescope and the Venus Radar Mapper mission.

The space station programme, which received official presidential approval last January, will not be able to purchase scientific instruments or support research and analysis activities with the \$8,000 million tentatively approved for the project. All of these expenses with presumably be borne by code E. NASA officials say that \$8,000 million will buy a basic manned module, two accompanying unmanned platforms, and an orbiting manoeuvrable vehicle to transfer payloads from the station orbit to higher orbits.

Some scientists, particularly those interested in Earth observation experiments, say that a single polar platform cannot be large or flexible enough to meet their needs, which means that the potential cost savings from sharing power supplies and communications equipment on the space station platforms will not be realized by the science budget. NASA appears to be resisting suggestions that a flotilla of small platforms be built in place of the large platform.

NASA has just sent out its request for

proposals for initial design studies of the station, and the specifications include the two large platforms. Contractors for the initial design work will be chosen on 1 April next year. (John Hodge of the space station programme told Banks's committee this summer that the large platforms provided a more "unified" approach to the programme and would be well suited for commercial projects, such as manufacturing in space.)

Despite such concern, space scientists are reluctant to criticize NASA's space station plans. The frequently heard argument is that, because the space station is going to be built in any event, the best thing scientists can now do is to offer their full support. In that way, they can hope to influence some design details to ensure that some good science can be carried out.

Burton Edelson, director of the Office of Space Science and Applications, told the Space and Earth Sciences Advisory Committee last week that, because his office had been "supportive" of the space station over the years, it had already been able to make the project "more suitable" for science. He said there would not have been unmanned platforms at all if it were not for his office's advocacy.

At the same time, some say that there are great political pressures on the NASA science programmes to make use of the space station, which is seen by many as long on engineering and short on purpose. The same reservations were expressed, to a lesser degree, about the role of science in the shuttle programme. The Office of Space Science and Applications is spending 13 per cent of its current budget on shuttle projects.

The Banks committee, significantly, was not asked to decide whether scientific

activities could better be carried on the station, but only what activities could be carried out on it. It seems agreed that the space station opens up some new possibilities but most of the experiments listed by Banks's committee, and by the Space Science Board in a separate study, could clearly be done as well if not better without the station.

Some relief for the science budget may come from abroad. NASA administrator James Beggs has been pushing hard for Japanese and European contributions to the cost, to the tune of \$1,000 million and \$2,000 million respectively. There is some talk of the foreign contributions going for a second polar platform.

Stephen Budiansky

ESR to Grenoble in Franco-German deal

Paris

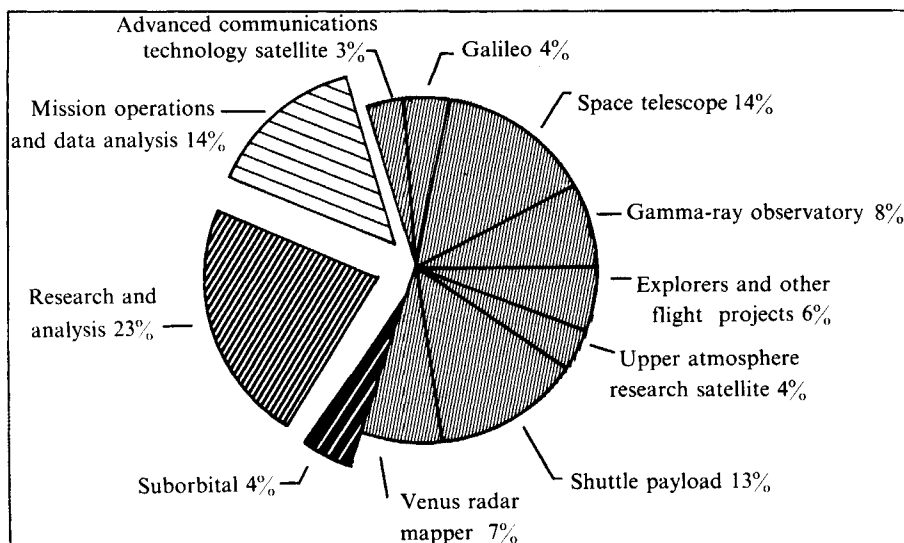
OFFICIALS representing Hubert Curien and Heinz Riesenhuber, the French and German science ministers, last week had to answer for the Franco-German decision to site the European Synchrotron Radiation Source (ESRS) at Grenoble.

On Friday in Brussels, they met their opposite numbers from the smaller European states under the auspices of the European Science Foundation (ESF) "progress committee" which — according at least to Denmark and Italy — had been gathering and evaluating site proposals. Instead, the committee was faced with a decision by France and Germany that ESRS should be built in Grenoble.

In fact, the first notice other governments had of the "decision" — except for press reports — came in a joint letter from Curien and Riesenhuber only on Tuesday last week, too late for official reactions by Friday. What may now emerge is an agreement between France, Germany, Britain and Italy to which the smaller states, notably those of Scandinavia, might adhere in a separate consortium. But Denmark, clearly the most hurt is unwilling to participate. The Netherlands is also playing cool.

Meanwhile, Britain, which France and Germany dearly want as a partner, while strongly backing the Grenoble site, has no money to offer, and may enter only with observer status "or something between that and full participation". The French certainly are not keen to see ESRS start up as a purely Franco-German institute, as the Institut Laue-Langevin was before 1967.

And why did France and Germany (and in effect the United Kingdom) ignore the ESF progress committee? Because it had too low a political level, say French sources, and would have been unable to trade one facility for another. This, finally, is what France and Germany did with ESRS and the trans-sonic wind tunnel, now to be built at Cologne. Robert Walgate



Europe in space

Italy treads Columbus path

US pleas for European collaboration in the planned space station have won hearts and minds in Italy. The Italian minister of research, Mr Luigi Granelli, was in London last Friday to persuade the faint-hearts of the British Government that they should contribute some £40–50 million a year from 1986. For Granelli, Europe's only future lies in space. "We've fallen behind in electronics and in biotechnology; we just can't afford to fall behind in space as well." His own government appears to agree. But Britain is still counting its pennies.

Mr Granelli's enthusiasm may yet work in sceptical London the magic it has worked in Rome. Earlier this year, seeking an increase in the budget of the "national space plan", Granelli told his cabinet colleagues that they should either cut the annual 50,000 million lira (£20 million) plan to nothing, or increase it four-fold. "It would have been a waste to continue spending just the 50,000 million lira" he said last week. The Italian cabinet seems to have agreed, to have pulled down their visors and taken a walk in space with Mr Granelli. From this summer, Italian space spending has risen to 150,000 million lira, and will reach 200,000 million lira next year (see box). Italy's contribution to the European Space Agency (ESA), at present 120,000 million lira, comes on top.

Safety in numbers

THE increased Italian space budget is mostly to be spent on technology, but X-ray astronomers may just benefit. The package includes approval for a "phase-B" detailed design of an X-ray satellite (SAX), Italy's first, due for launch at the end of the decade. Italian astronomers have been seeking outside collaborators on instruments for SAX which already, according to some observers, has too many small instruments incorporated to satisfy this or that hungry group of Italian astronomers.

International collaboration could stretch SAX further. Discussions are now at a critical stage with Dutch high-energy astronomers, but Britain and West Germany have been occupied with their collaboration in the German-designed ROSAT. Now, however, Italy may face a problem with the Dutch Government, which has set its face against further satellites after last year's infrared satellite IRAS.

Apart from SAX, other items on the Italian agenda include ITALSAT (a telecommunications satellite) IRIS (a Space Shuttle upper-stage that would put SAX in orbit), TSS (a tethered "satellite" to hang in the atmosphere 100 km beneath the Shuttle) and LAGEOS, a geodesy satellite (also launched by an IRIS-Shuttle combination).

Robert Walgate

Granelli urged the same argument last week on British industry and information technology minister Geoffrey Pattie. Granelli is convening the first ministerial meeting of ESA since 1977 in Rome next January to plan Europe's response to the US invitation. He hopes that British ministers will not be dragging their feet.

Europe needs "coherence" and "a vision" in the face of the National Aeronautics and Space Administration (NASA)'s plan to begin detailed design of the space station (phase B) on 1 April, 1985, and to start construction in January 1986. A series of bilateral arrangements with NASA would be dangerous, he believes, so he seeks a 70 per cent increase in the ESA budget in the period 1986–91.

The Italian vision, shared by West Germany (see *Nature* 25 October, p.696) is that of Columbus, and adjunct to the space station consisting of a manned laboratory (developed from Spacelab), an experiment platform and a control module. But Granelli, statesmanlike, says he is more interested in principles than details. There are as many technical proposals for collaboration as there are space technicians in Europe, he admits. Whether or not Europe builds Columbus matters less than that the old continent should safeguard "our standing and negotiating power with the United States".

In Britain, the issue appears to be whether Europe should plan a truly independent manned capability in space (as seems to be the view in Italy, West Germany and France), almost without counting the cost, or whether the approach should be more gentle. Nevertheless, it is stressed in London that "Britain wants to do something European". There are fears that mainland Europe is beginning to write off Britain as a potential partner. Collaboration in the space station may be a test.

The effect of all this on the space science programme is cloudy. Granelli said last week that he did not know what would happen to the recommendation that ESA's mandatory space science programme should increase by 40 per cent by 1990, mostly for astronomy.

In Britain, cabinet sources suggest that cash for any space station collaboration should come out of "a political pocket", leaving existing space sciences unaffected, but Granelli is not so sure. "When you talk of science, you must think beyond the existing community of space users." Granelli was thinking of "science in zero-gravity" — materials processing and so on. But he added that Earth scientists and astronomers now using space through ESA should not "in principle" suffer, although "someone has to make a choice in the end." So something may be sacrificed on the altar of the space station.

Robert Walgate

Anglo-French cooperation

New guise plea

FRANCE and Britain should set up a joint foundation with the aim of linking universities and industry, the French President, François Mitterrand, told the British Parliament last week. A new idea? No. But each time it pops up it takes a slightly different shape, and each time Britain knocks it down again.

The idea first surfaced in Edinburgh two years ago, during a meeting of the largely commercial Anglo-French Council, when the research director of the French oil and chemicals company Elf-Aquitaine proposed a joint research and development planning agency. This was to be heavily weighted towards industry, and was to devise more projects like the European Airbus.

That plan died until Robert Chabbal, a director at the ministry of research in Paris, transformed it into a proposal for a series of bilateral science and technology seminars. These have taken place — but between France and Sweden, and France and West Germany, rather than with Britain. The British view is that France has gained more than Sweden or West Germany from the meetings so far.

Meanwhile, the science counsellor at the French Embassy in London has been trying to convince his British colleagues that some kind of collaborative structure should be established between Britain and France, but without much success. The French view is that collaborations move best beneath a well-defined governmental "umbrella agreement". The anarchic British, however, say there are already plenty of routes for collaboration, and therefore no need for new structures.

President Mitterrand's intervention may have been stimulated by the suggestion by British science minister Peter Brooke, at the Council of Europe last month, that Europe should establish an international science academy. The French consider that this may be a sign of enlightenment in Britain, perhaps even a basis for their own idea of a science-industry forum.

The British are still reluctant, however, and suspicious that France merely wishes to tap British invention. They may be right. M. Guy Denielou, president of France's only technological university (Compiègne) said last week that he fully approved of an institution such as Mitterrand had proposed.

"Of course, I would hope to be a customer" he said. "British research is too big for British industry... All your ideas get taken up in America or Japan." It would be a good thing, said Denielou, if British science could feed some ideas into French industry as well. The same thing could also happen in reverse, he said. But the suspicious feeling in Britain, rightly or wrongly, is that the traffic would work out to be one-way.

Robert Walgate

Chinese nuclear power

Japan sets sights on business

Tokyo

THE chief obstacle to an atomic energy cooperation pact between China and Japan, the unwillingness of China to permit verification of non-military uses, seems likely to be removed in the coming weeks. China is now apparently prepared to allow some sort of inspection of nuclear plants supplied by Japan to be carried out under the supervision of the International Atomic Energy Agency (IAEA).

The exact form that inspection will take is not, however, clear. On the Japanese side, Michiyuki Isurugi, minister of state for the Science and Technology Agency, was willing only to say that Japan had received a "new proposal" in response to its wish, repeated over four rounds of talks, that China accept supervision by IAEA or provide other concrete means of guaranteeing peaceful use of atomic energy facilities. Isurugi did, however, express confidence that he would be able to travel to Beijing in the next few weeks to sign a pact. The Chinese side seems to have been rather more forthcoming, according to an article in the *Asahi Shimbun* claiming that officials in Beijing had stated they were now willing to allow IAEA inspection.

Earlier this year, Japan had received promises that China would allow "goodwill" visits of Japanese experts to Japanese-built plants in China. This was enough for Mitsubishi to go ahead, with government encouragement, and win a ¥1,500 million (\$6.3 million) contract, against strong French and West German competition, for pressure vessels (using US Westinghouse technology) for one of China's first nuclear reactors. But lack of a full cooperation agreement was thought by the nuclear industry to hinder its competition with the French, West German and British companies that have joined the rush to build nuclear power stations in China.

Much is at stake, for China is the last major nation without any nuclear power generating facilities and has committed itself to installing 10,000 megawatts of capacity — a minimum of ten reactors — by 2000. US companies, whose home market has dried up, should be the strongest in the field, but are for the time being sitting on the sidelines, waiting for the US-China nuclear cooperation agreement announced last April to be approved by (or at least submitted to) Congress (see *Nature* 309, 657; 1984). France, Italy and Brazil have already signed nuclear agreements with China — although none is so extensive as that signed by West Germany — and the Soviet Union has also indicated its willingness to sell reactors to China, apparently without assurances on strictly non-military use.

Meanwhile, construction has begun in China on two nuclear plants, one of 300 megawatts at Qinshan near Shanghai,

which will use a Chinese-designed pressurized water reactor (PWR) with pressure vessels supplied by Mitsubishi, and the other with two 900 MW PWRs near Canton, which will supply part of its power to (and be partially financed by) Hong Kong. Final contracts for the second reactor have not yet been awarded but discussion has reached an advanced stage with Framatome of France for the reactors and the General Electric Company of the United Kingdom for the turbines.

Two more power stations are in the offing: one near Shanghai and a second in the northern Liaoning province. If US activity remains frozen, contracts for one of these could go to West Germany, to judge from Chancellor Kohl's optimism during his recent visit to Beijing.

When the US presidential election is over, though, the United States-China atomic energy accord is expected to be given high priority. If US companies get

back into the field, competition will be much more intense. But Japanese manufacturers still fancy their chances, because they have the best operating efficiency record in the world and believe China will go for reliability.

US commentators, on the other hand, have quoted Chinese officials as saying that they want to learn from the "teacher, not the student" — virtually all of Japan's nuclear technology is derived from US designs. China is, however, almost certain to distribute its contracts among the advanced nations both to encourage cut-throat competition and to ensure that it can always ride out a sudden worsening of relations with one or more of its suppliers. Nor can the apparent willingness of China to allow inspection of Japanese plants be regarded as a signal that similar concessions will soon be made to the United States. Last April, when the possibility of "goodwill" inspection came up, the Chinese said they were taking into account Japan's unique position as the only country in the world to have been hit by nuclear weapons.

Alun Anderson

Atmospheric chemistry

Preparing for crises ahead

Washington

CITING "essential unanswered questions" about changes in the Earth's atmosphere, a US National Research Council panel has proposed a long-term international study of tropospheric chemistry, to be led by the United States. At a public meeting here, panel members agreed on the need to understand basic atmospheric processes in order to provide sound advice to policymakers and to avoid "crisis responses" to environmental problems. The panel's chairman, Robert Duce of the University of Rhode Island, said the proposed "Global Tropospheric Chemistry Program" could by the end of the century allow the impact of atmospheric changes to be anticipated.

Several panel members conveyed a sense of urgency about the unknown effect of documented atmospheric changes. Ralph Cicerone, the panel's vice chairman, said it was "shocking" how little is understood about cycling of atmospheric constituents. Dr Jerry Mahlman of the National Oceanic and Atmospheric Administration, an independent supporter of the programme, drew attention to increasing atmospheric levels of trace gases, especially methane, which could have an effect on temperature "equal to or greater than" carbon dioxide and which "can be expected to cause a 20°C cooling of the stratosphere within two or three human lifespans". Mahlman said nobody knows why methane levels are rising and that "we ought to find out".

The research council panel had been convened at the request of the National Science Foundation and the National Aeronautics and Space Administration.

The proposed study was generally welcomed, but there were some reservations about the apparent dominance of a supposedly international programme by the United States: the Global Tropospheric Chemistry Program is seen as a large component of the International Geosphere Biosphere Program recently approved in principle by the International Council of Scientific Unions (see *Nature* 4 October, p.402). Henning Rodhe, professor of chemical meteorology at the University of Stockholm, pleaded for a genuine international effort and reminded the largely US audience that "there is some atmospheric chemistry outside the United States".

The new programme would in its early stages concentrate on developing new instrumentation and classifying ecological habitats by their effect on the troposphere. Later research would focus on biological sources of tropospheric constituents, global distribution of trace gases, studies of photochemistry and investigations of removal processes. New global models would incorporate scavenging and chemical reactions in cloud droplets.

There has so far been no serious attempt at costing, but, according to one "very crude" estimate, the \$10-20 million spent annually on tropospheric chemistry in the United States would need to be doubled to get the programme started. Detailed proposals are to be drawn up at a series of workshops organized by Dr Cicerone. But the panel acknowledged the difficulty of persuading Congress to support a programme unlikely to produce useful results within 10 years.

Tim Beardsley

Cancer mortality

US half-rate goal in dispute

Washington

THE National Cancer Institute (NCI) continues to defend its "achievable" goal of halving US age-adjusted cancer mortality by the year 2000, first announced (to great dramatic effect) before the US Senate appropriations committee in April this year. But NCI has yet to make public any detailed justification for this claim, and its feasibility has been questioned. NCI plans to publish a full description of its year 2000 programme in 2 or 3 months, after review by the National Cancer Advisory Board.

Since 30 per cent of the 440,000 deaths from cancer each year in the United States are attributable to cigarette smoking, all agree that curtailing smoking is crucial. NCI's provisional estimate is that it should be able to halve lung cancer mortality by 2000 by preventive measures alone. But there are doubts whether NCI can achieve the 50 per cent reduction in smoking over the ten years to 1990 needed for this target.

As a government agency, NCI is prevented from lobbying for what is known to be one effective means of reducing smoking: higher taxes on cigarettes. NCI's weakness on this score was underlined in the summer, when Congress voted to lower cigarette taxes as part of a compromise that included strengthening the health warning on cigarette packets. For the same reason, NCI cannot take a position on, for example, banning cigarette advertising or restricting smoking in public places. And NCI's own advertising efforts are limited to free "public service" air time donated by radio and television stations and pamphlets distributed in physicians' waiting rooms. The emphasis is on helping only those who want to stop smoking, rather than encouraging confirmed smokers to reform.

Even if smoking could be halved by 1990, success in halving lung cancer by the end of the century would be far from assured. While smoking appears still to be decreasing among males, there are signs that it is increasing among females, and such reductions in smoking as there have been in recent years have been restricted to the higher socioeconomic classes. In order to reach the target, the NCI programme would have to be effective among the older/heavier smokers who are most at risk but the hardest to convert, according to Dr John Higginson of Universities Associated for Research and Education in Pathology. Dr Lewis Kuller, director of the department of epidemiology at the University of Pittsburgh, goes further, saying that the year 2000 goal on lung cancer will be achievable only if a chemopreventive agent is discovered for treatment of those who have stopped smoking.

The target is defended by Dr Joseph Cullen of NCI. The institute is about to embark on large-scale intervention trials

that will include 10 million subjects over 5 years, designed the better to define how to persuade people to stop smoking. There will be a special programme for high-risk groups, and physicians will be trained in anti-smoking counselling. Cullen predicts that NCI will extend its influence with policy-makers and the media, and that the continuing shift in public attitudes will be helped by, for example, boycotts against magazines that carry cigarette advertising. Cullen will not be drawn on how NCI will get around the statutory limits on its authority except to say that more federal and state funds will have to be used.

Smoking aside, NCI plans to achieve a further 15 per cent reduction in cancer mortality by the year 2000 by ensuring full public access to state-of-the-art treatment. Dr Peter Greenwald, director of NCI's division of cancer prevention and control, says this target makes no allowance for advances in treatment yet to be discovered or for earlier detection, but that it is based on treatments now at the stage of clinical trials. On top of the reduction due to improved access, Greenwald anticipates another 15 per cent reduction due to improved treatment of micrometastases; he points to experimental treatments using monoclonal antibodies targeted at specific oncogene products. While acknowledging that the figures might be disputed by a conservative statistician, Greenwald stresses the importance of having goals.

Others doubt whether these dramatic predictions can be defended. While more widespread screening programmes could markedly reduce breast cancer mortality (possibly by up to 30 per cent), and perhaps cervical cancer also, the other major killers seem stubbornly resistant to improved treatment. Kuller says he "would have difficulty being convinced" that these improvements are in prospect without further major advances in early detection. Tests for cancers of the gastrointestinal tract, for example, are still much needed.

Perhaps the most controversial element of NCI's year 2000 plan is the goal of reducing mortality by 5 per cent by changing eating habits. While inter-population comparisons indicate that up to 35 per cent of cancers may be caused by dietary factors, the precise components involved are not known. Despite the small number of studies of fibre intake and cancer, NCI's director, Dr Vincent DeVita, told the Senate Appropriations Committee in April that a doubling of fibre intake could reduce colon cancer by 30 per cent by the year 2000. NCI also now believes that a reducing fat consumption will help.

Others are not so sure: Kuller says the 5 per cent target "could be wishful thinking". In 1982, the National Academy of Sciences concluded that it was impossible to quantify the likely effect on

cancer incidence of dietary changes. Higginson believes the effect of diet on cancer incidence in adults has been exaggerated, and points out that the reductions in fat consumption seen in recent years, which have had a marked effect on heart disease, have so far failed to have the effect on cancer incidence that would be expected according to the simple dietary hypotheses, thus making the overall NCI target even less realistic.

Tim Beardsley

European environment

Hopes to reform farming

Brussels

THE European Community may be about to mount a vast new programme of research and development in agrochemicals. That, at least, is what will happen if the commissioner for environment and internal market affairs, Mr Karl-Heinz Narjes, gets his way. At a conference in Dublin two weeks ago, Mr Narjes argued that pesticides and fertilizers have had an undesirable effect on the environment, and that the time has now come to encourage technical innovation to yield more selective but less persistent pesticides.

The environment commissioner has been trying for some time to persuade his colleagues in the commission that environment policies should be integrated with others, and with agricultural policy in particular. Formally, this principle is written into the current action plan on the environment. Now there are signs that the member states are coming round to Mr Narjes's way of thinking.

Mr Narjes is not anti-farmer. Instead, his aim is that agriculture should continue to profit from the use of chemicals without putting an intolerable burden on the environment. In Dublin, he was pleading for more work on the toxicology of pesticides, allowance for regional and national differences of cultivation techniques and a training programme for farmers.

Within the Community, the Dutch Government has already advocated the use of European Community funds to protect the environment linked to farming in less-favoured areas. Meanwhile, the United Kingdom, backed by Ireland, has been urging that Community agricultural funds should be made available to farmers for "environmentally friendly actions".

The outlook for Community regulation of agricultural pollution seems less bright, however. Member states have been unable to agree on a list of acceptable pesticides since 1976, while proposals for maximum permitted levels of pesticides in foodstuffs have been gathering dust since 1981. Mr Narjes seems bent on circumventing the regulatory log-jam by tackling the problem at source.

Anna Lubinska

Soviet meteorology

Bolts from out of the blue

UP to 25 per cent of recorded weather information fails to reach the Soviet Union's Hydrometeorological Scientific Research Centre, the new director, Aleksandr Vasil'ev, has told an *Izvestiya* correspondent. The interview is also the first indication that Vasil'ev had been promoted from his former post as deputy, thus generating speculation that his predecessor, Dr Mikhail Petrosyants, had been dismissed for failing to forecast the hurricane which caused widespread devastation in the Russian Federal Republic on 9 June.

The storm, apparently caused by a warm dry airstream from the eastern Mediterranean meeting a cold airstream from the Barents Sea, has cost the Soviet Union upwards of 120 million rubles (\$150 million) in restoration work, compensation and insurance. Devastation of houses and "production facilities" is reported from Ivanovo, Gor'kii, Kalinin, Kostrona and Yaroslavl as well as in the Chuvash republic, with an unspecified number of deaths. A special resolution of the party central committee and the Council of Ministers to make good the damage done by "a natural calamity" was passed in mid-July.

The Ivanovo hurricane was followed almost immediately by three further meteorological calamities — extensive flooding in the Baikal area in mid-June, a hurricane in the Volynia district in Western Ukraine in mid-July and flooding in the Amur basin "complicated by Typhoon Holly" in late August. But in these cases, although there was considerable damage to agriculture and production schedules, some warning seems to have been given. In the case of Ivanovo, however, there was apparently no warning.

Reliable weather forecasting is particularly important to the Soviet planned agriculture, where the logistics of sowing and harvesting are determined regionally. Moreover, Soviet meteorologists still have to live down the memory of last autumn, when the unexpected onset of severe weather left more than 40 ships of the Arctic mercantile fleet stuck in the ice.

In the event, the new icebreaker *Leonid Brezhnev* extricated the vessels. The Soviet press at the time stressed the valuable lessons learned for future Arctic navigation, which may be why the northern navigation season has been ended several weeks earlier this year. But *Pravda* also noted a year ago that Arctic weather was not by itself "guilty" of causing the emergency, and that all those responsible for the ships' presence in the ice so late in the season should share the blame. Coming only a few months later, the Ivanovo hurricane may well have made it necessary for the meteorological service to offer a scapegoat. And, at 65, Petrosyants is already past the official retirement age.

But is Vasil'ev justified in blaming the problem on poor information flow? The question is pressing because of the new drive to extend the use of "improved" land, which has to be made fit for agriculture by either irrigation or drainage. Last week's central committee plenum admitted that all is not well with the much publicized food programme, launched by Leonid Brezhnev only in 1982, and "further measures" were urged by Mr Chernenko including a major extension of irrigation and the bringing in of virgin land, in Soviet Central Asia. River diversion projects, allowed to lapse during the Andropov era, seem on the way back. Yet even the most optimistic planners admit that such schemes can only yield the projected harvests if every drop of irrigation water is carefully husbanded. Failures in weather forecasting could have serious results.

The Soviet meteorological service

appears to have the technical means. It maintains a union-wide network of weather-watch stations, including "out-posts" such as that on Franz Josef Land, and a fleet of weather ships, which not only participate in forecasting programmes but also attempt to correlate Soviet weather conditions with such factors as the change of sea temperature in the northern Pacific. As early as 1970, experiments were under way with networks of automatic recording stations which would feed data automatically into a computer.

Reliable weather forecasting is also one of the main justifications for the Soviet space programme, while the export of meteorological equipment and the training of operators form an important part of Soviet aid programmes to developing countries. Vasil'ev's suggestion, that it is not the information gathering but the information transmission system which is at fault, is one to which the State Committee for Hydrometeorology and Environmental Control will no doubt pay attention.

Vera Rich

Corporate research

Monsanto goes for independence

St Louis, Missouri

THE Monsanto Chemical Company's new \$150 million Life Sciences Center in Chesterfield, St Louis, opened this week and will employ 1,200 scientists, engaged mainly in basic biological research, by 1987. About one third of the scientific budget of the company is earmarked for biotechnology research, and Richard Mahoney, Monsanto's chairman, is prepared to invest \$25–\$75 million before seeing a return in the marketplace, at least ten years from now.

The first commercially-produced biotechnological product will be a bovine growth hormone (methionyl bovine somatotropin, MBS) which will increase milk yields by 15–20 per cent, thus allowing farmers to reduce the sizes of their herds. But the blue-suited Monsanto managers are most proud of their company's commitment to basic research, in which they claim now to be self-sufficient. The basic biology of the MBS development had to be contracted out to Genentech.

Monsanto will, however, continue supporting university research programmes, including the 5-year \$23 million collaboration with Philip Needleman's group at Washington University (St Louis). Here, the discovery in the mammalian heart of atrial peptides affecting sodium balance and smooth muscle tone (*Science* 223, 67; 1984) has led to current investigations of the role of these peptides in human blood pressure control.

Rooftop greenhouses and 123 "growth chambers" at the new institute are the two most visible signs of the company's investment in plant genetics and physiology. In plants, expression of genes via the soil

bacterium *Agrobacterium tumefaciens* may eventually confer characteristics such as insect or disease resistance and increased seed production; now Monsanto is extending its earlier work on pea plants (*Science* 223, 838; 1984) to petunia and tomato. The commercially attractive goal of expressing genes in monocotyledonous plants is, however, well advanced in European laboratories; altered *A. tumefaciens* plasmids can be expressed in asparagus (see *Nature* 311, 594; 1984) and plants of the Liliaceae and Amaryllidaceae families (*Nature* 311, 763; 1984).

At last week's opening ceremony, Frank Press, president of the National Academy of Sciences, endorsed Monsanto's research investment, but warned of the dangers of an imbalance between federal and industrial research spending. (Industry now spends more on basic research.)

Press attacked the government practice of supporting work by a single investigator in a single department, and stressed the need for a different basis for the US grant system. Monsanto is ideally placed to make the massive financial investment needed for a multidisciplinary attack on scientific problems and says it intends to continue publishing its research promptly in the scientific literature.

It claims to have learned from past mistakes in the arena of public opinion, being eager to comply with the National Institutes of Health guidelines for recombinant DNA research and preaching the benefits of biotechnology research to the public. So long as it remains open about its work in this highly competitive area, this huge boost to basic biological research will be welcome.

Maxine Clarke

Spanish universities

Reforms beginning to work

Barcelona

SPANISH universities are in a state of upheaval. The effects of last year's Law for University Reform are making necessary important changes in the structure of universities, affecting in particular the teaching staff, the number of departments, the statutes of the universities and their control.

Just before the end of the academic year, steps were taken to reduce the number of professors under contract and to give some of them tenure (see *Nature* 308, 395; 1984). Commissions were appointed for each discipline, which then decided who qualified for permanent positions. The results now published show that around 6,000 out of 8,000 candidates qualified, a proportion that the Ministry of Education believes accurately reflects the situation.

Mount Wilson Inc.

Washington

A CALIFORNIA industrialist has formed a corporation to take over the Mount Wilson observatory. The present owner of Mount Wilson, the Carnegie Institution of Washington, which also operates the Las Campanas observatory in Chile, announced this summer (see *Nature* 15 August, p.88) that it can no longer afford to keep both sites and will shut down Mount Wilson on 1 July 1985 unless a new owner can be found.

The new corporation, called the Mount Wilson Research Corporation, was founded by George Roberts, who made his fortune designing and manufacturing aircraft lavatories. His company, INCA Corporation, also made the zero-gravity lavatory for the space shuttle.

The corporation has no funds as yet, but hopes to raise an endowment of several million dollars, earnings on which will provide the basic annual operating costs of the observatory. Foundations and individuals are being approached for contributions.

According to Dr Arthur Vaughan, one of several research astronomers associated with Mount Wilson who are involved in Roberts's effort, the corporation is talking to universities and organizations such as the Smithsonian Institution and the Jet Propulsion Laboratory about entering into a consortium.

Vaughan said that the aim is both to maintain an in-house research staff at Mount Wilson and to serve the needs of outside researchers. To that end, all principal investigators would be involved in managing the use of the facility. The board of directors is to be made up of non-astronomers who will be concerned largely with the financial health of the corporation.

Stephen Budiansky

The results varied considerably between the different commissions, and although there was no obvious difference in the quality of the candidates, acceptance rates varied between around 50 per cent and more than 80 per cent. Those not selected are preparing to do political and legal battle. There will be more positions next year, and the situation of university professors will have changed profoundly in a period of only two years. Some people fear the long-term effects, particularly because the number of students is beginning to decrease and there is a danger of saturation of university positions soon.

The statutes of universities will have to be adapted to the new regulations. All universities have been asked to make proposals, and these will have to be approved by the government. The body in charge of this task is the General Assembly (Claustro) which includes professors, students and administrative staff. In some universities, conflicts have arisen between diverging interests.

An important change in the structure of universities, which will be reflected in the new statutes, is the definition of university departments. There is widespread agreement that the present departments are too small and too numerous, having mostly been built around a single professor. It is hoped that the new regulation will stimulate the concentration of people, administration and equipment in order to build up more efficient structures.

From now on, departments should have a minimum number of professors (twelve in the latest version of the law), and a full-time professor as the head of department. Normally, there should be only one department in any subject in a single university. The names of the departments should in principle correspond to areas of knowledge defined by the ministry. In some cases, these regulations may cause problems, for instance in medicine and engineering faculties where many departments have mostly part-time professors or in departments where the members are physically separated or working in very different research subjects. In any case, the power of full professors ("Catedráticos") will probably be diluted in these bigger departments.

The most important change is that the national government is transferring control of the universities to the autonomous governments of the different regions of Spain. Agreement to hand over the universities to the Government of Catalonia has already been reached. But strong differences still exist about the amount of money to be transferred and, as a consequence, the process is at a standstill. The government hopes that this period of upheaval will result in more dynamic universities.

Pedro Puigdomènech

Star wars killings?

Washington

ALTHOUGH almost all of the \$1,300 million budgeted this year for the Strategic Defense Initiative (star wars) is committed to research projects that have been under way for some time, defence contractors are already sensing a kill. "The Department of Defense has budgeted \$24 billion over the next five years for the Strategic Defense Initiative", reads an announcement from the American Society of Mechanical Engineers (ASME). "Will you find a place in this state-of-the-art cornucopia — expected to be a greater undertaking than the Apollo program?" To help them out, ASME is sponsoring a day-long seminar in Washington in mid-November, featuring briefings from key defence officials to show "how to get in on the ground floor". The admission fee is \$245 a head.

The only new funds to come out of the star wars programme so far are \$6 million to be awarded for an initial "creative effort to develop methodologies for defining alternate architectures". Proposals are due by 1 November; it is expected that more than one contractor will be chosen. The winners will be asked to "develop the strategic defense mission", considering both "treaty constrained and unconstrained threats", and to study various "candidate" anti-ballistic missile technologies.

Stephen Budiansky

Balzan prizes



PROFESSOR Jan Hendrik Oort of Leiden University in The Netherlands (above) has won this year's Italian Balzan Prize for astrophysics. The genetics prize has been awarded to Sewall Wright (95), professor emeritus in the Universities of Chicago and Wisconsin. The winner of the prize for history and criticism of the literatures is Professor Jean Starobinski of the University of Geneva.

Each of the winners will receive 250,000 Swiss francs. In 1985, the three prizes will be awarded for mathematics, palaeontology and history of occidental art. □

North Sea pollution

West Germany calls for action

Brussels

WEST Germany faces other North Sea coastal states in Bremen this week with demands to make headway in the next two years on marine pollution and improved monitoring of the ecosystem of the valuable fishing grounds of the North Sea.

Participating in the International Conference on the Protection of the North Sea are ministers from the coastal states (Belgium, Denmark, France, the Netherlands, Norway, Sweden and the United Kingdom), from whom West Germany hopes to extract a political commitment to ratify and strengthen existing international conventions. The European Economic Community is also invited.

The North Sea is at present covered not by one but by several international conventions governing the protection of the marine environment from pollution from land-based sources (Oslo, Paris), from pollution by oil (Bonn) and from dumping (London). Not all have been ratified.

The West Germans stress the danger of latent chronic pollution (accumulation of chemicals) as well as more directly visible pollution such as oil. They are particularly concerned about the ecological sensitivity of coastal waters, river estuaries and mudflats, important feeding grounds for fish and birds.

The discharge of dangerous substances through inland and coastal waters is controlled by the 1974 Paris Convention. In 1976, the European Community also adopted a framework directive on emission standards for industrial waste discharges, but so far only limits on mercury, cadmium, polychlorinated biphenyl wastes, waste oils and lindane (last June) have been adopted, out of a list of 129 dangerous substances.

According to Greenpeace, the environmental pressure group, 1,120 tonnes of cadmium and 1,000 tonnes of mercury continue to flow annually into the North Sea. Twenty-three million tonnes of toxic sludge end up in Rotterdam harbour.

Some 11 million tonnes of largely untreated sewage sludge also end up in the North Sea, with the result that algal growth in the North Sea has increased considerably over the past 5–10 years, according to the West Germans.

The West German proposals also include a system for licensing wastes, discharges, and international information exchange on the quantities discharged and requirements that the pollution load should be reduced. At the beginning of this year, West Germany introduced a domestic system of fines for pollution of the North Sea.

According to the West Germans, there is a definite need for further action, including better enforcement and an extension of the Marpol conventions. They also propose that the North Sea be designated as a

special area, that sea dumping should be reduced and eventually dropped. The European Community is preparing a directive banning the dumping of black-listed substances and limiting others.

To combat oil pollution, West Germany proposes a region-wide system of airborne surveillance, improved cooperation in port state control, more use of deep-sea pilots' services and greater compliance by drilling rigs with rules the oil content of effluent.

West Germany also wants to extend research and development programmes to include more biological monitoring and the sampling of more substances. The Paris Commission (the secretariat of the Paris Convention) already monitors mercury, cadmium, polychlorinated biphenyls, oil

from refineries and reception facilities and drilling platforms as well as radioactive effluent from processing plants. The Community intends to start monitoring titanium dioxide wastes next year. West Germany suggests that all this information should be collected in an international database.

Community support for the West German proposals is important, if only because Community directives are binding on member states, whereas conventions enter into force only when ratified. It is therefore unfortunate that, in the case of the draft directive on accidental marine spills of oil and other dangerous substances, now under discussion, West Germany, together with Belgium, Denmark, the United Kingdom and France, feels that a Community directive is neither "necessary nor desirable".

Anna Lubinska

No canker cure from China

Washington

US DEPARTMENT OF Agriculture (USDA) scientists working to combat the citrus canker now rampant in Florida (see *Nature* 27 September, p.293) are frustrated by a ban on new agricultural science exchanges with China. The US scientists were hoping to use a symposium planned for this month in China as an opportunity to collect new citrus germ plasm, possibly including canker-resistant strains. But the symposium has now been called off, and researchers are wondering whether the ban will be lifted in time to plan a visit for 1985.

The ban was ordered a year ago by the under secretary of agriculture for interna-

tional affairs, Daniel Amstutz, in retaliation for a Chinese decision to cut imports of US grain, breaking a contract. At the time it seemed likely that China would suffer more from a ban than the United States. The Chinese action was in turn a retaliation for US limits on textile imports.

Southwest China is believed to be the region where both citrus plants and citrus canker originated, according to Dr Jack Hearn of USDA's horticultural research laboratory at Orlando, Florida. American and Chinese scientists last year exchanged plant materials of interest, but citrus canker had not then appeared in the United States.

Tim Beardsley

Whaling

Japan pleads to carry on

Tokyo

THE Japanese Government is to try again to persuade the United States not to retaliate if Japan breaks the whaling quotas set by the International Whaling Commission (IWC) earlier this year (see *Nature* 18 October, p.595). Talks will be held in Washington on 1 and 2 November between the director-general of Japan's Fisheries Agency, Hiroya Sano, and John Byrne, director of the US Department of Commerce's National Oceanic and Atmospheric Administration. Earlier attempts to reach agreement ended in failure on 16 October.

As a result, the Japanese sperm whale coastal fishing fleet (owned by the Nihon Hoge and Nitto Hoge companies) put to sea on Friday 19 October. By the following Wednesday, however, before they had had a chance to catch a whale, the four boats were ordered to return home to give new talks a chance of success.

IWC set this year's sperm whale catch quota at zero, in the absence of agreement among the member countries. Japan,

which had been expecting a minimum quota of 400, vital to keeping its coastal industry going, thus considers it has been unfairly victimized by pressure from conservationists elsewhere.

The Antarctic Ocean minke whale quota was also cut, this time on a scientific recommendation, to a level that would give Japan a share amounting to 1,941 whales — some 600 fewer than the number at which Nippon Kyodo Hoge, Japan's only far-seas whaling company, says it can sustain commercial operations. The company is not yet abandoning whaling, however, and its mother ship, the *Dai San Nisshin Maru*, set sail from Yokosuka for the Antarctic on Saturday 27 October. It hopes that an agreement can be worked out with the United States before it reaches its catch quota around next January.

If the United States stands firm, then the Packwood-Magnuson amendment will be invoked if Japanese boats break the IWC rulings. Japan's fishing quota in the United States 200-mile zone could then be cut to zero.

Alun Anderson

Seascale hazard

SIR — The letter from Pomiankowski (*Nature* 13 September, p.100) raises two points which he considers introduce doubts about the conclusions of the report of the Black Advisory Group into the possible increased incidence of cancer in West Cumbria.

The first concerns the appropriate index on which to rank the occurrence of lymphoid malignancy in the 675 electoral wards of the Northern Children's Cancer Registry region. The suggestion is that the report was incorrect to include a table showing the wards with the highest incidence rates per 1,000 children, and that using instead the associated Poisson probabilities would have been more appropriate. It is important to recognize that these two sets of figures are not simple alternatives, since they measure different aspects of the disease frequency. The rate per 1,000 is a direct measure of the proportion of children under 15 years of age registered with lymphoid malignancy over the period 1968–82, which gives 9.73 per 1,000 in Seascale — calculated on the basis of a total of four cases during the 15 years and the 1981 census population. The cumulative Poisson probability, on the other hand, is a measure of the chance of finding the observed rate (or an even greater one) if the true (rather than the estimated) local ward rate were the same as the overall regional rate of 0.61 per 1,000.

The first index is an assessment of the average risk of a child in the area developing lymphoid malignancy before 15 years of age, whereas the latter is an indication of how unlikely it is that the raised rate in Seascale could have occurred by chance under a particular hypothesis. Each of these indices has its role and importance. In the case of a small-sized population and a small number of cases, the observed rate is subject to sampling fluctuation — and the associated Poisson probability helps to assess the size of this. Both figures are therefore informative and are given in the report. It is not true, however, to refer to the latter as representing the "incidence" of lymphoid malignancy — as was done by Mr Pomiankowski.

The second point concerns two childhood cases of lymphoma in Seascale registered since 1982. It is said that the exclusion of these from the calculations of rates and Poisson probabilities is "disturbing". The figures given in the report were provided by Dr Craft and his colleagues from the Northern Children's Cancer Registry for the years 1968–82. At the time (late 1983/early 1984) when these figures were being prepared, it was too early to include comprehensive cancer registrations beyond 1982 over the whole Northern Region since there is inevitably some delay in the notification of cancer cases to the regional register and, thus, in the availability of complete annual records for inclusion in analyses. For this reason it

is premature to consider the effect of these two cases — known about because of the special attention to the village — on the ranking of Seascale: it remains to be seen whether or not further cases of lymphoid malignancy have moved other small communities up into higher rankings, either of rates or probabilities. No doubt Dr Craft and his colleagues will do this in due course, and make further assessment of the situation possible.

M.J. GARDNER

*MRC Environmental Epidemiology Unit,
(University of Southampton),
Southampton General Hospital,
Southampton SO9 4XY, UK*

Defining length

SIR — On 21 October 1983, the Comité International des Poids et Mesures (CIPM) promulgated a new definition of the metre as: the length of the path travelled by light in vacuum during a time interval of $1/299\,792\,458$ of a second¹. This new definition raises some fundamental and practical problems.

At the fundamental level, one must reformulate Einstein's second relativity postulate: "Light is always propagated in empty space with a definite velocity which is independent of the state of motion of the emitting body"². For Einstein, this was an empirical fact. Now, this has become a definition and the corresponding empirical fact is that the length of a solid body depends neither on its spatial orientation, nor on the inertial frame where that body is at rest. (For example, in Michelson's celebrated experiment, the length of the arms of the interferometer did not change as that instrument was rotated or as it was carried by the Earth in an orbit around the Sun.) It is easy to derive the Lorentz transformation law from the above postulate.

On the practical side, this new way of defining length implies that there must be two *synchronized* clocks at the extremities of the measured object, in order to determine the corresponding time interval. As synchronization is best performed by light signals², this effectively means that in order to measure the length of an object, one ought to observe the round trip time of a light signal and divide the result in two.

Moreover, the return trip of the light signal must follow precisely the path along which the signal was sent. This is because a closed path surrounding a nonvanishing area (with a suitable set of mirrors) yields a result depending on the direction of travel, if the body being measured rotates. The difference in travel time (clockwise versus counterclockwise) is³

$$t = 4\omega A/c^2,$$

where ω is the angular velocity and A is the surrounded area.

If this effect is ignored (the round trip rule is not even mentioned in the CIPM recommendations!) the circumference of the Earth at the Equator is 123.9m longer if

measured eastward than if measured westward.

ASHER PERES

*Department of Physics,
Technion-Israel Institute of Technology,
32 000 Haifa, Israel*

1. Giacomo, P. *Am. J. Phys.* **52**, 607-613 (1984).
2. Einstein, A. *Ann. Physik* **17**, 891-921 (1905).
3. Landau, L. & Lifshitz, E. *The Classical Theory of Fields*, 282 (Addison-Wesley, Cambridge, Mass., 1951).

Red Sirius

SIR — The recent correspondence (*Nature* 6 September, p.8) concerning the colour of Sirius reminds me that in the ancient astronomy of Iceland, the star was called Loki's Brand, obviously because of its periodic disappearance into the "underworld", but also because of its flaming colour. This apparent colour-naming is therefore not just a Greek phenomenon, but something which doubtless has considerable significance.

NIGEL PENNICK

*142, Pheasant Rise,
Bar Hill, Cambridge CB3 8SD, UK*

Mount Wilson costs

SIR — Your report (*Nature* 15 August, p.88) on the Carnegie Institution's plans to sell the Mount Wilson Observatory understates the scientific importance of this facility and overstates its operating costs. The \$1.56 million budget for Mount Wilson listed in the Carnegie Financial Report — apparently your source for this figure — includes all astronomers' salaries, fellows' salaries and support operations both for Mount Wilson and for the Las Campanas Observatory that Carnegie operates in Chile. A California consortium hoping to operate Mount Wilson asked the director to estimate operating cost for Mount Wilson alone. A detailed estimate, made by our business manager, was produced: \$0.75 million per year. This includes the fraction of all engineering and support services spent on Mount Wilson, but of course not those devoted to Chile.

Many individual astronomers and a number of organizations within and outside the United States have inquired about obtaining all or part of Mount Wilson Observatory. The reason is that it is an excellent site for stellar and solar astronomy. Your implication that this work is less exciting than extragalactic astronomy simply reflects a bias, one that is certainly not shared by all astronomers.

At the recent (October 1984) European meeting on solar physics, the current Mount Wilson solar and stellar research was described as vital, and a resolution was adopted urging that all possibilities for its continuation be explored.

DOUGLAS DUNCAN

*Mount Wilson and
Las Campanas Observatories,
813 Santa Barbara Street,
Pasadena, California 91101, USA*

Can Europe avoid the Space Station?

from Erhard Keppler

The European nations could easily be tempted into a very expensive alliance with the United States to produce a space station. But is it worth it?

RECENTLY, there have been talks between the US space agency, NASA, and various European governments in which the American side has aimed to reach agreement with European countries to share in the proposed US Space Station project, in particular its costs.

The space station is supposed to be a large facility in a near-equatorial Earth orbit, about 300 km above the surface. It could be composed of many smaller units launched separately but docked together in space. Such units may, according to NASA's futuristic view, reach dimensions of hundreds of metres in length. One possible configuration of such a station is a long tube, bent into the circumference of a wheel with spokes towards a centre for stabilization and finally put into slow rotation to provide centrifugal forces in the tube to substitute for the gravitational force on the Earth's surface.

On this view, a space station would have a crew ranging from 10 persons at the beginning to perhaps 100 some years later. The first station would, according to NASA plans, become operational in about 1992, and could (again according to NASA statements) serve either as an operational base for launches of Shuttle-launched spacecraft into other orbits, as an Earth observation station or possibly as the first stage towards a production facility in space. These are the essentials of the arguments put forward in support of NASA's space station, the costs of which are established to be about \$11,000 million in the next eight years.

Because of the uncertainties attending such projects, these cost estimates may easily grow to remarkable numbers, as has been shown during the development of the Space Shuttle. NASA's offer to the Europeans is then to participate in the development and manufacture of a space station within a cost-envelope of several thousand million dollars.

The offer has put Europeans in a difficult position. First, Europeans, who had their first experience of manned spaceflight only recently in the first Spacelab flight, are unable as yet to make a sound judgement as to the benefits of this endeavour especially as the second opportunity for a Spacelab flight is not due until 1985. In any case, large numbers of flights are ruled out by the tremendous costs per flight, now more than DM100 million (£25 million). At the same time, it is to be assumed that the US side will not support an eventually successful European effort in, say materials

science by providing them further support, frequent flight-opportunities, for example.

Second, Europeans are afraid that the United States may eventually achieve a major success by using their space station commercially (for satellite launches, production and so on) so giving US industry better conditions by, for example low charges, so that European companies might be put at a competitive disadvantage. Thus it may turn out that two-step launches (from ground to a space station and from a space station into geostationary orbit) are much cheaper than direct ascents. Then European space companies would have no chance to compete with US companies in the vital communications market. But such apprehensions have no real basis; they are games in sandboxes.

Third, the conglomerate Europe has so far not been able to find a common strategy in space research and space technology. European space efforts are rather the sum of some countries' space efforts or bilateral agreements.

The case against

What should Europeans do in this situation? What are they able to do? The following are some of the conditions that must be dealt with, and the consequences which must be considered.

- One remarkable argument for the US Space Station, analogous to European apprehensions, is based on the irrational fear that, because the Soviets may build a space-station and so gain advantages denied to the United States, the United States should also build one just in case.

- The Space Science Board of the US National Academy of Science has rejected NASA's plans for the space station, which military sources consider unduly vulnerable.

- Many European scientists reject the space station because it is scientifically uninteresting and because they are afraid that the high costs associated and with its development and operation will draw away resources which are urgently required for more reasonable research.

So what have we gained so far from manned spaceflight? What is the prospect of a firm and solid basis on which to move further in this particular direction? Interestingly enough, the most frequently heard word in this context is "probably", as in its frequent assertion that the materials sciences are probable customers

of a space station, as well as the biological/medical sciences and, last of all, also the Earth observing sciences.

Astrophysics is, however, readily excluded because the required accuracy of observations, demanding constancy of viewing direction, is unattainable from a space station, which will always suffer from vibration due to the motion of the astronauts, from impacts of dust and small particles of meteoritic origin and from thermal tensions and relief. Astrophysics as well as Earth-observation techniques can, however, live perfectly well with unmanned spacecraft, specially because such missions should preferably extend over many years and be carried out in eccentric or polar orbits that will always be forbidden to manned missions because of the radiation load that astronauts would suffer.

Materials science is said to be another candidate customer. Materials scientists have just started to investigate the effects of reduced gravity on various rather elementary processes, which are certainly of high scientific interest. But at this early stage, there is not even the faintest hint that it will be necessary to have production facilities in space. Semiconductors, for example, can be made sufficiently well on the ground; neither crystal imperfections due to growth under gravity nor contamination have so far prevented the industrial production of microprocessors or large memories. There are other limits, of course, but there are also other interesting technical possibilities which may become the starting point for new technological pathways that are perfectly well pursued on the ground. Studies under zero gravity may in any case be satisfactorily continued on rocket flights or from unmanned satellites. The recovery of probes from such spacecraft is a well-proven technique, widely used in military applications. There is no justification whatsoever to mount a costly new programme on such evidence.

Finally, the interest of biology in zero-gravity conditions has greatly lost its attraction in the past few years. Indeed, space medicine flourishes largely because there are astronauts.

Much of the enthusiasm for the manned space station as a zero-gravity laboratory seems to stem from ignorance of what has already been achieved. The new "space science discipline" of materials science is carried on by scientists who have never worked with space technology. They often

do not know about the ingenious robots designed over the years by hundreds of scientists to investigate stars, planets, our Sun or the Earth's atmosphere. Automatic self-controlled sequences (such as opening an oven, removing a melted sample, replacing it by another, closing the oven and repeating the procedure) do not have to be invented afresh. Such simple tasks have been possible for many years.

A physical experiment provides answers to the question for which it was designed. If the answer is negative, another experiment will be set up in order further to attack the problem. The rule is that other questions have to be asked after certain answers have been acquired. It is a fiction to believe that such experiments require the presence of people. If the question asked by an experiment is clever enough, the answers will be "yes" or "no". To design such an experimental set-up so that it will not be disturbed by external measures is a challenge to the skill and imagination of the experimenter.

Passengers

The very presence of astronauts in a space station requires extraordinary efforts. More than half of the cost of the development of Spacelab could have been saved if the systems had been unmanned. But for manned systems, the effort will always be greater than is really necessary, for nobody will take responsibility for endangering people's lives. Efforts of this kind are concerned with the number of subsystems and their redundancy, with their reliability, with the selection of materials (beginning with their outgassing performance) and with endless functional tests, the complexity of which is a nonlinear function of their number. The very size of the system requires an enormous amount of testing; the Spacelab Manual, for example, is several thousands of pages thick.

The requirements of reliability, and the selection techniques and materials, result in extremely high costs for all on-board systems, including instrumentation. But parts ordered to the necessary high standards of reliability have prices higher by factors of 100 than those of commercially available parts. Many are only available in the United States, where they are produced mostly for military purposes. In contrast, unmanned satellites can now be built at relatively lower cost, taking advantage of the improved reliability of modern electronic circuits, even those manufactured commercially. Indeed, the use of the Space Shuttle, a manned transporting system, as well as Spacelab, a manned facility, has in fact inhibited developments that would have led to lower costs. Experiments flown on Spacelab, if they had flown on unmanned satellites, could have cost a fraction of what was paid for them.

But unmanned spacecraft can be launched by Ariane, the European launcher, the development of which would be without logic if Europe moved towards

the Space Station, which will require the exclusive use of the Space Shuttle.

Moreover, European participation in the US Space Station would require European industry to use the expensive electronic parts made in the United States. As for the major part of the job, the life support system, European companies would have to build their share under American licences. The chief developments for European industry in such a participation could well be the production of screws and the bending of sheet-metal into interesting shapes.

In contrast, the development of unmanned space stations would force Europe to become more independent from the United States and could bring about the development of important technologies that are at present not available in Europe, perhaps not anywhere. This does not mean that joint efforts with the United States should not be sought. In the past, they have been fruitful, in many cases. But we should not move towards such cooperation on a space station.

There is no problem in constructing spacecraft in modular principles so that subsystems can be removed from the outside. Such repair can be performed by robots as easily as by men in space. Indeed, robots would be able to perform such repairs in any possible orbit — polar, eccentric or geostationary. There is no argument in favour of manned operations from the "repair" point of view.

Indeed, a recent study at Massachusetts Institute of Technology has recognized that robots, controlled according to anthropomorphic principles, may be the most promising route to future space techniques. This development was started 20 years ago when K. Kleinwächter in West Germany built such a system (called "Syntelmann"). There would have to be commodious communications links to send commands to a robot in space, supposedly to perform some operation. The robot may, for example, have a simulated arm. On the ground, a person would do the same experiment, when sensors monitoring the motion of the person's arm would relay appropriate commands to the robot and cause it to do exactly the same thing: opening a box, moving something, turning it. Such devices are also required on the ground, for operations in toxic environments, for example.

Opportunity

With this starting-point it would be tempting to develop, for example, unmanned spacecraft to support zero-gravity research. All the investigations required at this stage of the research could be accomplished in such a way. It would be necessary to develop sensors to generate images in all parts of the electromagnetic spectrum, from X rays to the far infrared. So as not to increase bit rates necessary in the communication channel, techniques

such as on-board image evaluation, pattern recognition, redundancy reduction in information and the improvement of error detection and error-correcting codes for information transfer would be needed. If Europe, recognizing the enormous potential generally of such developments, were to move into such a direction, this would not only stimulate a broad field of scientific research (including for example, artificial intelligence by the application of the concept of associate memories) but would also require from industry products of high technological standard. The result would be to push European science and technology far ahead of where they are now — far beyond what participation in a space station programme might yield.

So why should Europe move along the dead-end route of the US Space Station? All experts agree that it would indeed be a dead-end. It is true that the idea has generated much science fiction — colonies in orbit, on the Moon or on other planets. But there is no possibility that we will be able to live on any other planet in the Solar System, or on an asteroid or on the Moon. There is no air to breathe while gravity is less than on the Earth, with possible consequences for human physiology. Neither the Moon, nor Venus or Mars have strong magnetic fields, so that their surfaces are bombarded by the full intensity of the cosmic radiation. The Moon has no atmosphere, that of Mars is dilute, so ultraviolet radiation and X-rays penetrate to the ground at high intensity. These two environments are a terrible threat to life. Venus has dense layers of ice clouds (from sulphuric acid) in its atmosphere, and almost 100 bar pressure at the surface, so it is dark and hot there; life would require permanent protection. Mercury is also extremely hot, and the outer planets are dark, as on Earth after sunset. So the dreams of life on other planets are unreal, pure fiction. We will stay on this planet. Why not accept that?

Is it not possible for a continent with the tradition of Europe, with its ability and with more than 300 million people, to stand on its own? Why not break with the traditional policy of imitating what is being done in the United States and, instead, move jointly in a promising direction which has been chosen deliberately by argument, not just by dreams or fears? If political unification is difficult in Europe, it should at least be possible to pursue common technical interests. Why not start with a powerful European space programme, designed to make good Europe's technological backlog, that will eventually generate contributions of general importance to our technical and technological development? Participation in the space station holds no such promise. So let us not participate. □

Erhard Keppler is at the Max-Planck-Institut für Aeronomie, Postfach 20, D-3411 Katlenburg-Lindorf, FRG.

Recollection of a simpler time

The death of Paul Dirac a week ago is a reminder of the precocious and staggering achievement, sixty years ago, of one of the most diffident of contemporary heroes

THE death of P.A.M. Dirac, in Tallahassee, Florida, on 20 October is a moving reminder of what science was like when everything was much simpler than now. The tale is that, as a shy but clever son of a Swiss-born schoolteacher in Bristol in the West of England, he embarked on a degree course in electrical engineering at the local university, just two months before the First World War ended. With a degree but no job, he took a degree in mathematics, as it were to fill in time. By 1923, with the help of a scholarship and a government grant, he was a research student at the University of Cambridge. And then, two years later, at the age of 23, he had put Heisenberg's quantum mechanics, not then published, on an axiomatic foundation that still endures.

How could so much have been done in so short a time? Luck must be a part of any explanation. Dirac was lucky that his Cambridge supervisor was R.H. Fowler, then full of excitement about the statistical mechanics of stellar atmospheres. Dirac was also lucky that Heisenberg went to Cambridge to talk at the Kapitza Club (then Cambridge's private *phi beta kappa*) in July 1925, and sent Fowler a set of the proofs of his first paper on new quantum theory by the middle of August. Who to ask to read them but Dirac? The real luck, of course, was that the time was ripe for a radical upheaval. Planck quanta were real enough but the old quantum theory of atoms, which worked well enough for atomic hydrogen, was increasingly at odds with spectroscopy.

But Dirac's special contribution to quantum mechanics was its independence of what others were about. Heisenberg's first paper (*Zeits.f.Phys.* 33, 879-893; 1925), the proofs of which Dirac had seen, had set the agenda — how to generalize Bohr's prescription that quantum states are those calculated by classical mechanics which happen to satisfy a quantum condition pulled like a rabbit from a hat. Heisenberg's innovation was to describe the kinematics and the dynamics of quantum variables strictly analogous to the classical dynamical variables — position, momentum and so on.

Dirac's first paper on the subject was received at the Royal Society on 7 November 1925, and published the following month in *Proc. Roy. Soc. A* 109, 642-653 (which says something else about the simplicity of those times). Its contemporary importance was threefold.

It emphasized that the quantities representing dynamical variables in quantum mechanics are elements in an algebra then undefined, it demonstrated that the circumstance that two variables, say q (for position) and p (for conjugate momentum), do not commute is not an embarrassment (as it had seemed to Heisenberg) but a natural way of making a bridge with the quantities in classical Hamiltonian mechanics called Poisson brackets.

Dirac's achievement in the following five years was staggering. By the time his PhD thesis had been submitted (in the spring of 1926), he had made a first attempt at the problem of relativistic quantum mechanics. This was the train of argument that led to Dirac's relativistic equation of the electron which provided both a natural, not *ad hoc*, way of accommodating electron spin and as clear a prediction that positrons must also exist as any more recent prediction of a particle of matter. Almost in passing, Dirac quantized Maxwell's equations, providing the foundations of a quantum theory of radiation and a model for what is now called field theory.

Throughout this revolution, Dirac's style was distinctive. He was not so much an iconoclast as an adventurous spirit, sceptical of fudges but prepared to see where logic would lead. What else can account for his failure to accept as his own the idiom in which his elders and presumptive betters described the problems facing physics in 1925?

The point is nicely illustrated by Dirac's own algebra for the dynamical variables of a quantum system. Heisenberg had shown that the quantum condition boils down to the simple relation that $qp - pq = i\hbar$ if q and p are, respectively a position and its conjugate momentum (with i the square root of -1 and $\hbar$ Planck's constant divided by 2π). When Born and Jordan had reminded Heisenberg that his quantum equations embodied the rules of matrix multiplication, the three went on to elaborate their doctrine of matrix mechanics.

Not Dirac. The quantum variables were the elements of an algebra whose rules were dictated by the physics. Dirac called them q -numbers (*Proc. Roy. Soc. A* 110, 561-569; 1926) and said this about them:

At present one can form no picture of what a q -number is like. One cannot say that one q -number is greater or less than another. All one knows about q -numbers is

that if z_1 and z_2 are two q -numbers... there exist the numbers $z_1 + z_2$, $z_1 z_2$ and $z_2 z_1$ which will in general be q -numbers... One knows nothing about the way the numbers are formed except that they satisfy all the ordinary laws of algebra, excluding the commutative law of multiplication.

At any point along the way, Dirac could have resolved this uncertainty by changing idiom. Representing q -numbers by matrices would have given Heisenberg's quantum mechanics, representing them by differential operators would have given Schrödinger's wave mechanics. But why needlessly particularize?

The scale of Dirac's achievement during this short spell is now not easily comprehended. It is as if he had found a secret window on the foundations of physics, a companion for that through which Einstein looked in 1905. It cannot often have been that a physics textbook (*The Principles of Quantum Mechanics*, Oxford, 1930) should be nothing but a record of its author's research.

Dirac's distinctive style stamps this great book and the generation of students brought up on it — or in its shadow. Dirac himself, however strongly convinced in 1930 that the problems of physics had been solved, seems not to have been filled with proselytizing zeal. Always shy, he seems to have been tragically shy of students. Much of the 1930s were spent in a hunt for some way of ridding calculations of radiation fields of the infinities that bedevilled them. Dirac did not succeed in this herculean task, completed only in 1948 by Schwinger, Tomonaga and Feynmann.

But from this period there survive three novel ideas — the notions that magnetic monopoles may exist, that it cannot be a mere coincidence that dimensionless numbers such as the ratio of electrical to gravitational forces are equal to the age of the Universe (in atomic units) and (as a consequence) that the gravitational constant should decrease with increased age of the Universe (*Nature* 139, 323; 1937). Time, no doubt, will show what happens to these conjectures.

Dirac's style persists. The anonymous reviewer of the first edition of his book (*Nature* 127, 699; 1931) said of Dirac that "he bids us throw aside preconceived ideas regarding the nature of phenomena... we may describe this as the application of pure thought to physics". He was a subversive whose pure thought was more penetrating than most other people's. **John Maddox**

Molecular medicine

Therapy by instant evolution

from Robin Carrell

EVOLUTION has modified and refined the structure and function of proteins. An understanding of how this has happened makes feasible the prospect of artificially inducing similar changes to create specifically designed variants of human proteins. As a starter, on page 77 of this issue¹, Rosenberg *et al.* (Chiron Research Laboratories) report the use of recombinant DNA technology and yeast to replace the amino acid at the reactive centre of human α_1 -antitrypsin, a natural inhibitor of protease enzymes. The mutant α_1 -antitrypsin retains the specificity of the natural inhibitor, but promises to provide more effective protection to tissues during inflammatory stress. Of particular interest is its potential to prevent the development of the degenerative lung disease, emphysema, but the independent production of related mutants of α_1 -antitrypsin by two other genetic engineering companies, Transgene, Strasbourg (M.A. Courtney *et al.* *Nature*, in the press) and Biogen, Cambridge, Massachusetts (D. Schindler, personal communication) attests to much wider therapeutic implications.

This focus of activity, follows from the realization² that α_1 -antitrypsin is one of a family of serine protease inhibitors, the serpins, each of whose function is critically determined by a single amino acid at the reactive centre. These protease inhibitors have specific roles in the control of the proteolytic cascades responsible for immune reactions, for coagulation, and for other responses to inflammation. Since it is the aberrant activation of these cascades that underlies many disease processes, there is great medical interest in the possibility of specific intervention by the use of engineered inhibitors.

α_1 -Antitrypsin is the archetype³ of the serpin family which includes the plasma proteins: antithrombin, α_1 -antichymotrypsin, C₁-inhibitor and α_2 -antiplasmin⁴. The structural homologies of the family, together with recent crystallographic studies⁵ of α_1 -antitrypsin, indicate that the serpins all share the same highly-ordered globular structure. A feature of this structure is the exposure of the reactive centre as a loop on the exterior of the molecule. This loop is believed to fix the reactive centre in a conformation that makes it an ideal substrate for the target protease. The inhibitor and its target protease thus form a tight 1:1 complex which can be rapidly recognized and removed from the circulation.

The specificity of each inhibitor is critically dependent on a putative cleavage site at the P₁ residue (see the figure) of the reactive centre of the inhibitor. This was initially indicated by comparing the reac-

tive centre of α_1 -antitrypsin, which has methionine at P₁ and is primarily an inhibitor of elastase, with that of antithrombin, an inhibitor of thrombin which has arginine at P₁. Unequivocal evidence that the P₁ residue was responsible for the specificity came from an unusual experiment of nature⁶: a child from Pittsburgh was found to have a bleeding disorder associated with a new genetic variant of α_1 -

Serpin	P ₂	P ₁	P ₁ '	P ₂ '	P ₃ '	P ₄ '	Target
α_1 -Antitrypsin	Pro-MET-Ser-Ile-Pro-Pro-						Elastase
Valine mutant	Pro-VAL-Ser-Ile-Pro-Pro-						Elastase
Pittsburgh mutant	Pro-ARG-Ser-Ile-Pro-Pro-						Thrombin
Antithrombin	Gly-ARG-Ser-Leu-Asn-Pro-						Thrombin
α_1 -Antichymotrypsin	Leu-LEU-Ser-Ala-Leu-Val						Chymase
Mouse contrapsin	Arg-LYS-Ala-Ile-Leu-Pro-						'Trypsin'
C ₁ -inhibitor	Ala-ARG-Thr-Leu-Leu-Val						Kallikrein

Reactive site sequences of natural human serine proteinase inhibitors (serpins) and two mutants of α_1 -antitrypsin. Cleavage putatively occurs at the P₁-P₁' bond with a specificity of inhibition that is clearly related to the P₁ residue (C₁-inhibitor sequence is a personal communication from J. Travis).

antitrypsin, in which the reactive centre methionine had been replaced by an arginine; as predicted, the mutant α_1 -antitrypsin had lost the ability to inhibit elastase but had become a highly effective inhibitor of thrombin.

The general applicability of this dependence on the P₁ residue was recently demonstrated by Hill and colleagues⁷, from a comparison of the amino acid sequences deduced from the cDNA of mouse and human serpins. Their most remarkable conclusion concerned human α_1 -antichymotrypsinogen and mouse contrapsin. These two proteins have a similar sequence but the human inhibitor has a leucine at P₁ and hence inhibits chymotryptic-like proteases, whereas the mouse inhibitor, with a P₁ lysine, has no antichymotryptic activity but predictably is as an efficient inhibitor of trypsin.

Evolution, in several experiments of her own, has thus demonstrated how to prepare specific inhibitors. Rosenberg *et al.*¹ have applied these lessons to the production of a mutant α_1 -antitrypsin in which the P₁ methionine has been replaced by a valine. Their choice follows from the fact that about one in a thousand northern Europeans has a severe genetic deficiency of α_1 -antitrypsin and is consequently susceptible to cumulative damage to the elastic tissue of the lungs, which results in emphysema. In principle, the defences of susceptible individuals could be improved by intermittent transfusions of α_1 -antitrypsin. Genetic engineers, anticipating this need, have taken the opportunity to produce a modified form of α_1 -antitrypsin that should inhibit elastase

and so provide improved protection.

The risk of damage to the lungs in both normal individuals and those deficient in α_1 -antitrypsin is heightened by the susceptibility of α_1 -antitrypsin to inactivation by oxidants released by leukocytes or present in cigarette smoke⁴. There is strong evidence that this is due to the oxidation of the P₁ methionine to methionine sulphoxide, with consequent loss of effectiveness as an inhibitor of leukocyte elastase. Improved protection might be provided by the replacement of the P₁ methionine of α_1 -antitrypsin with a residue resistant to oxidation^{3,4} — valine seems a logical choice⁸ for an inhibitor of leukocyte elastase. *In vitro* tests of the mutant that fits that bill support those predictions. Before the material is ready for therapeutic use, several questions will need an answer. Does it matter that the mutant protein, as produced by yeast, does not have oligosaccharide sidechains? Antigenicity will probably not be a problem, but how important is it that the mutant inhibitor is likely to be less soluble and long-lived than the natural form? This will be more of a problem if the material is to be administered by transfusions than if it can be directed to the lungs as an aerosol. A greater problem will lie in the assessment and marketing of an agent whose protective effect, with regards to emphysema, is measurable over years rather than days.

Other related mutants that are at the testing stage may have more immediate medical applications. It can be predicted that the P₁ arginine mutant^{1,6} will be an effective medium-term anticoagulant for individuals with either a genetic or functional deficiency of antithrombin. Here, a shortened half-life due to lack of glycosylation may prove an advantage by increasing precision in the day-by-day control of anticoagulation.

In a wider sense, the new papers underline the rapid growth in the practical applications of molecular biology. When we suggested the possible production of the valine α_1 -antitrypsin³ it was done so almost tongue-in-cheek. Yet 18 months later it has been achieved. The speed with which molecular biology is growing has turned protein chemistry into an adventure rather than threatening it, and has converted the academic exercises of correlating protein structure with function into matters of direct, practical importance. □

1. Rosenberg, S. Barr, P.J., Najarin, R.C. & Hallewell, R.A. *Nature* 312, 77 (1984).
2. Hunt, L.T. & Dayhoff, M.O. *Biochem. biophys. Res. Commun.* 95, 864 (1980).
3. Carrell, R.W. *et al.* *Nature* 298, 239 (1982).
4. Travis, J. & Salveson, G.S. *A. Rev. Biochem.* 52, 655 (1983).
5. Löbermann, H., Tukuoka, R., Deisenhofer, J. & Huber, R. *J. molec. Biol.* 255, 4053 (1984).
6. Owen, M.C., Brennan, S.O., Lewis, J.H. & Carrell, R.W. *N. Engl. J. Med.* 309, 694 (1983).
7. Hill, R.E., Shaw, P.H., Boyd, P.A., Baumann, H. & Hastie, N.D. *Nature* 311, 175 (1984).
8. Nakajima, K., Powers, J.C., Ashe, B.M. & Zimmerman, J. *J. biol. Chem.* 254, 4027 (1975).

Robin Carrell, of the University of Otago Clinical School of Medicine, Christchurch, New Zealand, is currently Commonwealth Fellow at St John's College, Cambridge, UK.

Oceanography

Tides, solitons and nutrients

from Melbourne G. Briscoe

TIDAL energy mixes the coastal waters, providing nutrients for biological processes. The mechanism is indirect; the topography of the shelf-break, the stratification of the coastal waters and the tides interact to produce groups of solitons — strong, short, high frequency internal waves — and it is the breaking of these solitons as they propagate inshore to even shallower water that is responsible for the mixing. That is the suggestion offered by H. Sandstrom and J.A. Elliott of the Bedford Institute of Oceanography in Canada¹. Their idea, based on their work on the Scotian Shelf, seems sound but perhaps does not go far enough.

Solitons are so-called because of their similarity to the mathematically precise idea of a non-linear wave that retains its identity and form as it propagates along as a singular bump on the sea surface. However, in most situations an internal wave actually travels as a singular wave of depression, since there is usually a deeper layer of water underneath the pycnocline (the layer where the water density changes most quickly with increasing depth) than above it. As the internal depression on the pycnocline propagates, it tends to disperse into the free-wave components that compose it. This tendency is balanced by the non-linearities that arise from the finite amplitude of the depression and allow the wave to retain its soliton form. This fragile balance of forces is easily realized in equations, can sometimes be simulated in the laboratory, but probably never really occurs in the ocean; I therefore use the word soliton with some reservations.

Solitons are thought to be produced by the surface tide running up against the edge of the continental shelf. The pycnocline overlying the shelf is disturbed vertically by the tidal flows and a depression in the pycnocline begins to travel shoreward. There is some uncertainty about the mechanism that forms the depression: Sandstrom and Elliott consider that the surface tide forms an essentially linear, internal tide at the shelf edge, and that the leading edge of the propagating internal tide is the depression that finally resolves into the soliton¹. The scales of these processes differ greatly. The surface and internal tides (called barotropic and baroclinic tides, respectively, in the oceanographic literature) occur once or twice a day, depending on location, but the wave sequence in the soliton is of just a few minutes period. The horizontal scales of these three processes decrease from thousands of kilometres, to tens of kilometres, to hundreds of metres.

Because of the tidal generation of solitons they can be recorded every 12.4 h in those locations where the semi-diurnal tide

is dominant. One might also expect a spring-neap cycle in the intensities of the process, but I am aware of no good data that confirm this; since the dynamics of the generation of the solitons and of their eventual propagation depend on the strength and depth of the pycnocline, the inevitable changes in the hydrography with space and time probably obscure spring-neap effects.

Surprisingly, solitons are often easy to observe in the ocean. The interaction of wave currents in the soliton with short surface wind waves causes a modulation of the roughness of the sea surface that is visible to the eye and to radar. Landsat satellite images described by C. Sawyer² show that the entire eastern seaboard of North America is covered with a pattern of surface streaks characteristic of solitons. This is especially so in the 50–150 m depth range of the shelf, and is most marked in summer and early autumn, when the thermocline is well developed. (Georges Bank and its surroundings perhaps have the most solitons in the western North Atlantic; does this account for the nutrient supply and the fishing industry on Georges Bank?)

Sandstrom and Elliott propose that the surface tides at the shelf edge are converted to internal tides and then to solitons that dissipate on the shelf¹. They calculate a figure of 6×10^7 J per metre of crest length for their solitons. If extrapolated to the coastlines of the world this yields 1×10^{15} J, less than a third of a percent of the total internal tidal energy, a negligible contribution to the problem of how the energy is

dissipated. On the other hand 6×10^7 J dissipated on a 10 km wide shelf in half a day provides about 50 mW m^{-2} towards the vertical mixing of the water and the erosion of the base of the mixed layer. This is at least ten times the energy supply estimated from internal wave dissipation in the deep sea and several times larger than estimates of the energy available from typical winds.

A.R. Osborne and T.L. Burch³ have described solitons in the Andaman Sea near Sumatra that are 15 times more energetic than the examples of Sandstrom and Elliott. Even if all of the coastlines of the world were covered with such solitons, and all these solitons were dissipated in half a day, this would still only account for 0.3×10^{12} W, just 20 percent of what needs to be dissipated from the surface tide in shallow seas according to G. Miller's calculations⁴. (These take into account surface tide dissipation due to conversion to internal tides or solitons on the shelves for some areas of the world, but not for the North American east coast where so many solitons have been observed.)

So even this extreme estimate of the energetics and dissipation of solitons, we are not much closer to understanding where tidal energy goes, but Sandstrom and Elliott's suggestion¹ of the importance of solitons for shelf-mixing and for the supply of nutrients to the euphotic zone seems plausible on the Scotian Shelf. It may even be applicable to thousands of kilometres of other soliton-covered shelves. □

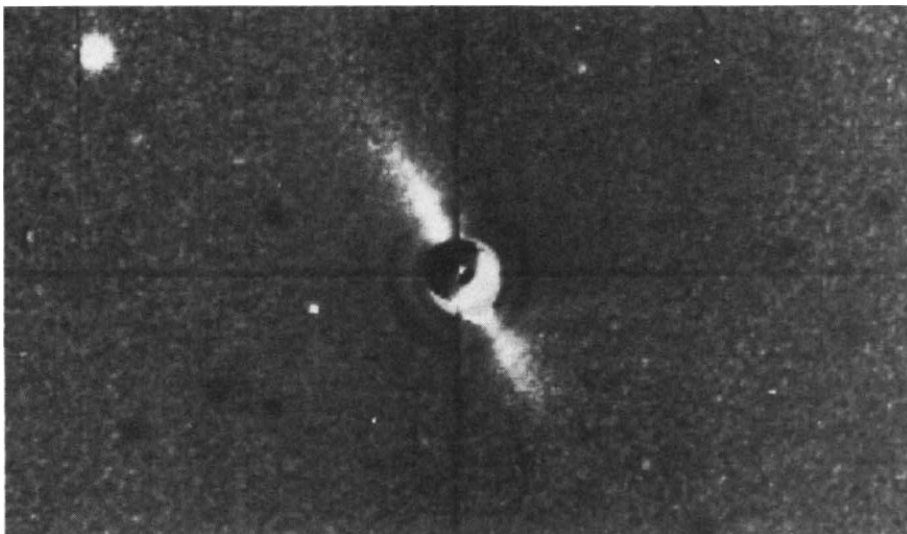
1. Sandstrom, H. & Elliott, J.A. *J. geophys. Res.* **89**, 6415 (1984).

2. Sawyer, C. *NOAA Tech. Memo. Envir. Res. Lab. Pacific mar. envir. Lab. Publ.* **46** (1983).

3. Osborne, A.R. & Burch, T.L. *Science* **208**, 451 (1980).

4. Miller, G. *J. geophys. Res.* **71**, 2485 (1966).

Melbourne G. Briscoe is at the Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.



A masked image of the star β Pictoris taken by R.J. Terrile (Jet Propulsion Laboratory) and B.A. Smith (University of Arizona) using a charge coupled device on the 2.5 m telescope at Las Campanas Observatory, Chile. The bright extended emission is thought to come from a circumstellar disk of silicates, ices and carbonaceous organic material. The disk, seen edge-on, is 40 billion miles across; the star is 50 light years away. This disk is the first of its kind to be seen clearly in astronomical photographs. (JPL and U. Arizona).

T-cell antigen receptor

The capture of the snark

from Miranda Robertson

WHEN I last discussed antigen recognition by T cells five weeks ago¹, I confined myself to what could be learned from the genes encoding the beta chain of the receptor heterodimer, on the grounds that there was serious doubt about the authenticity of the putative alpha-chain gene isolated by Saito *et al.*². There is no longer any doubt.

In this issue Hannum *et al.*³ report amino-acid sequences from a human T-cell receptor alpha chain that clearly have no homology with the amino-acid sequence predicted from the mouse gene of Saito *et al.*²; moreover, the alpha-chain protein is glycosylated whereas the predicted product of the gene has no glycosylation sites. And at the same time, Chien *et al.*⁴ and Saito *et al.*⁵ announce the isolation from mouse T cells of what both cautiously describe as a third type of T-cell receptor gene: in fact, thanks to the protein sequence of Hannum *et al.*, these genes can be unambiguously identified as alpha-chain genes. Yet the original Saito *et al.* gene², which I shall call the non-alpha gene, has all the important properties expected of a T-cell receptor gene; which leaves us, apparently, with one more receptor gene than the protein data had led us to believe we needed.

Since this surprising finding has given rise to speculation that antigen recognition in T cells may require two receptors, I shall briefly rehearse the arguments for and against such a schema in the context of what is now known about the expression of the alpha- and beta-chain genes, before going on to discuss a paper that contains a detailed analysis of several beta-chain genes and may bear on the alternative possibility of dual recognition by one receptor⁶.

How many receptors?

The notion that T cells may require two antigen receptors originated with the discovery that, unlike B cells, which recognize antigen on its own, T cells recognize antigen only in association with cell surface molecules encoded by the major histocompatibility complex (MHC). Specifically, cytotoxic T cells recognize antigen in association with class I MHC molecules whereas the regulatory T cells — the helper and suppressor cells — respond to antigen in the context of class II MHC molecules. Hence the suggestion that T cells might express two receptors, one for antigen and the other for the products of the MHC.

Under such a two-receptor scheme, the non-alpha gene might encode the MHC receptor. Its relative invariance, which was one of the strongest arguments against it as the alpha-chain gene, would be somewhat less of a handicap to a gene encoding a receptor for MHC molecules. Even so,

MHC polymorphism might demand rather more than the junctional diversity generated during the assembly of the separate gene segments encoding the variable (V) region of the receptor (for fuller explanation see ref. 1).

There are, however, more serious difficulties with this interpretation of the non-alpha gene. First, immunoprecipitation of the T-cell receptor brings down only the alpha and beta chains, together with the so-called T3 marker molecule, which is coordinately expressed on all T cells with the alpha and beta chains⁷ and whose major chain, unlike the non-alpha chain, has no homology with immunoglobulins⁸. Second, Chien *et al.*⁴ report that the non-alpha chain is expressed at much lower levels in T cells of thymus and spleen than the alpha- and beta-chain genes. Most serious of all, there is compelling evidence from cellular immunology^{9,10} that a single receptor is responsible for the dual recognition of antigen and MHC molecules.

It remains possible that the non-alpha chain exists as a minor isotype, analogous to the lambda light chain of immunoglobulin — which is also relatively invariant. In that case, however, it ought not to be expressed along with both the other chains in the cytotoxic cell line of Saito *et al.* (B cells expressing lambda light chains do not also express kappa chains). There is, in fact, so far no evidence that the non-alpha product is actually made and expressed on the surface of the cell, and since there are numerous precedents for aberrant transcripts the isotype hypothesis cannot be ruled out. On the other hand, until there is evidence that the non-alpha gene is expressed on the surface of one cell or another, its status as a functional gene must be considered unproven.

How like immunoglobulin?

How then is dual recognition reflected in the structure of the alpha-beta heterodimer? Before considering the structure of the beta-chain V regions, I must first dispose of the question of whether the differential expression of different alpha- and beta-chain gene pools could account for the differential recognition of class I and class II molecules by cytotoxic and regulatory T cells. I have already quoted¹ the evidence that the same beta-chain gene pool is expressed in cytotoxic and helper T cells, and it is now clear that both also express the same alpha-chain genes: the constant-region (C) segment of the alpha cDNA clone isolated by Chien *et al.* from a mouse helper cell is identical to that of the alpha-chain clone isolated by Saito *et al.* from a mouse cytotoxic cell. However, the variable regions are clearly very different,

which is to be expected in the light of the evidence for V-region diversity in the alpha-chain protein^{11,12}.

This brings us to the question of how the V regions of the alpha and beta chains might provide a structural basis for dual recognition. The homologies, both in sequence and in organization, between immunoglobulin and T-cell receptor genes, have already been discussed in *News and Views*^{1,13}. Because immunoglobulin recognizes only antigen, however, while the T-cell receptor recognizes both antigen and MHC, one would also expect profound differences. Such differences, chiefly in the pattern of diversity of the V regions, are now reported by Patten *et al.*⁶.

I have already mentioned¹ two ways in which beta-chain V-region diversity seems to differ from that of immunoglobulins: first, there is so far no evidence for somatic mutation in beta-chain genes; and second, there seems to be more scope for diversification during the assembly of the three gene segments encoding the variable region, in that it is possible in principle to assemble V genes either with no D segment or with more than one (for details and explanatory diagrams see ref. 1). What Patten *et al.* now show is that any such mechanisms for increasing somatic diversity of the receptor genes are acting on an already very heterogeneous set of germline V segments.

They have analysed beta-chain V sequences from seven helper cells and find that, unlike the immunoglobulin V segments, which tend to run to family sizes of up to 30, V_β gene families seem not to exceed two or three sequences. Moreover, there is very low homology between families: the least homologous immunoglobulin V segments are about as similar as the most homologous V_β sequences. More detailed analysis of this diversity has shown that it can be attributed to seven hypervariable regions, three equivalent to the classical hypervariable regions of immunoglobulin, two more which also have counterparts in immunoglobulin, and two that are exclusive to the T-cell receptor. From predictions of secondary structure based on primary sequence data, and on the assumption that the tertiary structure will be similar to that of immunoglobulin, Patten *et al.* conclude that the three classical hypervariable regions would, like those of immunoglobulin, form the antigen-binding pocket of the molecule, but giving a deeper cleft; while the remaining hypervariable regions would fall on the exterior of the folded protein and might mediate interactions with MHC molecules.

They further speculate that MHC binding might induce allosteric changes in the antigen-binding pocket that would represent a further source of diversity in the range of binding combinations possible. Some very preliminary unpublished evidence is quoted by Patten *et al.* in support of such allosteric interactions: two T cells recognizing the same antigen in association

with different MHC molecules express different V-region genes.

What next?

At this stage, there are not enough data to justify detailed speculation on the mechanism of dual recognition. If the non-alpha chain plays a part in the recognition of MHC molecules, then one might expect that, unlike the alpha- and beta-chain genes, it would be differentially expressed (if it is expressed at all) in different subsets of T cells. Tonegawa has preliminary evidence suggesting that it may indeed be transcribed only in cytotoxic and not in helper cells. In that case, one would also predict an equivalent non-alpha, non-beta chain mediating the recognition of class II MHC molecules in the regulatory subsets.

With all the relevant genes now isolated — both the genes encoding the MHC mole-

cules that the T-cell receptors recognize, and the genes encoding the receptor that does the recognizing — it should be possible to devise transfection experiments that will answer the questions raised in this article. □

1. Robertson, M. *Nature* **311**, 305 (1984).
2. Saito, H. *et al.* *Nature* **309**, 757 (1984).
3. Hannum, C.H., Kappler, J.W., Trowbridge, I.S., Marrack, P. & Freed, J. *Nature* **312**, 65 (1984).
4. Chien, Y. *et al.* *Nature* **312**, 31 (1984).
5. Saito, H., Kranz, D.M., Eisen, H.N. & Tonegawa, S. *Nature* **312**, 36 (1984).
6. Patten, P. *et al.* *Nature* **312**, 40 (1984).
7. Weiss, A. & Stobo, J. *J. exp. Med.* (in the press).
8. Terhorst, C. *et al.* *Nature* (in the press).
9. Kappler, J.W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198 (1981).
10. Hedrick, S.M. *et al.* *Cell* **30**, 141 (1982).
11. Kappler, J., Kubo, R., Haskins, K., White, J. & Marrack, P. *Cell* **34**, 727 (1983).
12. McIntyre, B.W. & Allison, J.P. *Cell* **34**, 739 (1984).
13. Williams, A. *Nature* **308**, 108 (1984).

Miranda Robertson is Acting Life Sciences Editor of Nature.

Space missions

Very-long-baseline interferometry takes to the sky

from P.N. Wilkinson

RADIO interferometers have long since surpassed the best angular resolution achievable with ground-based optical telescopes but radio astronomers have continued to strive for higher resolution in the quest of new information about cosmic objects. Since 1967, when the technique of very-long-baseline interferometry (VLBI) was introduced, we have been aware that there is a natural limit to this process — the separation of our telescopes is restricted by the size of the Earth itself. Nonetheless, there have so far been enough scientific rewards from exploiting existing radio telescopes for us not to worry too much about limited resolution.

Now, however, in common with many other astronomers, our thoughts are moving beyond the Earth towards an observatory in space, QUASAT (for QUASar SATellite), which would be linked to the earthbound network of radio telescopes. The initial ideas for this project were discussed recently at a workshop* sponsored by the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA), attended by astronomers, engineers and administrators from twelve nations.

Much of the present interest in VLBI stems from the recent extension of the principle of 'aperture synthesis' to arrays that span intercontinental distances; as a result, radio maps with milli- and sub-milliarc second resolution can now be made routinely. But many extragalactic sources and components in water-line masers are barely resolved even on the longest Earth baselines and at the shortest wavelength (1.35 cm, a frequency of 22

GHz) at which most large radio telescopes work well. Higher resolution, at least on continuum sources, is achievable at millimetre wavelengths but millimetre-wave VLBI is only in its infancy. To take advantage of the experience at centimetre wavelengths and to use the more numerous and much larger centimetre wave telescopes, the physical length of the baselines must be extended. This can be done only with the aid of a radio telescope in space.

QUASAT's astrophysical importance depends on its sensitivity, and on the resolution and quality of the synthesized image. In this 'first step' mission compromises must be made between these criteria. Thus the present proposal, which is dominated by the desire to observe at 22 GHz, has a relatively small antenna and hence sensitivity at the lower frequencies has been conceded for resolution. Nonetheless, QUASAT would have plenty to look at in its up to 5 year lifetime. The proposed orbit would yield roughly a three-fold better linear resolution, or a 10-fold smaller pixel area in the resulting image than Earth-based arrays. At 22 GHz the resolution is about 100 μ arc s, corresponding to a size scale ranging from 1 AU (the distance from the earth to the Sun) at the distance of the galactic centre (10 kpc) to about 10,000 AU at a distance of 100 Mpc. Computer simulations presented by A. Readhead (Caltech) show that with QUASAT and typical 10-station ground arrays, it would be possible to make radio images of a quality similar to that now achieved with multiple 'snapshot' observations using the Very Large Array (VLA), but with almost 1,000 times higher resolution; and that VLBI work in the Southern Hemisphere would be revo-

lutionized. Furthermore, using a worldwide network of 34 ground observatories which could be available in the 1990s, it would be possible to produce radio images of the compact nuclei of active galaxies as complex as any that the VLA has so far produced.

There was no shortage of exciting ideas for what the QUASAT mission could achieve. Superluminal sources — in which there are ordered motions with apparently faster-than-light speeds — are among the obvious targets. Even after a decade of observations, we have no real physical understanding of these phenomena. A closely related goal is to study the propagation of the energy-carrying beams, which emanate from the 'central engine' in active galactic nuclei, as they pass through the region that is responsible for producing the broad optical emission lines. We now know that in the much larger narrow-line region, the line-emitting gas clouds can bend and brake the beams; thus the interaction probes the physical conditions both in the clouds and in the beams. QUASAT will extend this study to the broad-line region, which is inferred to be less than a light year across in many cases, and which therefore cannot be studied directly even with the Space Telescope.

On the smallest accessible scales QUASAT could test the long-standing prediction that the brightness temperature of incoherent synchrotron emission is limited to 10^{12} K by inverse Compton scattering. It cannot be tested using Earth baselines but the QUASAT baselines would record 10^{13} K emission if it exists. The super-bright emission could be from a previously unsuspected coherent radiation source or, more likely, it could indicate that the emission region is approaching us at relativistic speed. In this case the apparent brightness temperature in the observer's frame is increased because of the effects of aberration.

Several speakers stressed QUASAT's great potential for studying the physical conditions in hydroxyl (at 1.6 GHz) and water (at 22 GHz) masers which are associated with star-forming regions and some late-type stars. Perhaps the most exciting prospect, however, is to use the water-masers to make fundamental distance measurements. The internal motions of water-maser spots, already revealed by Earth-based VLBI, have enabled the distances of three galactic maser complexes to be measured by statistical parallax techniques (Moran, J.M. *Nature* **310**, 270; 1984). QUASAT would greatly facilitate these measurements and would allow the distances to many more star-forming regions to be determined. And M.J. Reid (CFA) introduced the even more tantalizing prospect of using masers to measure the distances of nearby galaxies, thus helping to clarify the contentious extragalactic distance scale. Powerful water masers are known in some nearby

* Grossenzersdorf, Vienna 18–22 July 1984.

galaxies (Moran, *op. cit.* and Claussen, M.J. *et al. Nature* **310**, 298; 1984) and we could use the statistical parallax technique if we could measure their internal motions with an accuracy of a few microarc seconds per year.

With QUASAT, it might also be possible to determine the orbital motions of maser sources around a galaxy. These angular motions (corresponding to velocities of a few hundred km s⁻¹) will be an order of magnitude larger than the internal motions in the masers themselves. If the rotation curve and the inclination of the galaxy are known, then a distance estimate can be derived directly.

One debate at the workshop centred on the scientific trade-offs between raw resolution and imaging capability. A vocal minority took the view that much cruder images, but with two or three times greater resolution, might lead to more fundamental discoveries. Where to draw the line between resolution and image quality is a matter of judgement but clearly more computer simulations are needed to investigate the point at which image quality deteriorates to levels unacceptable to all.

R.Z. Sagdeev (Space Research Institute, Moscow) gave a brief, but up-to-the-minute account of current Soviet thinking. The USSR hopes to place a large (30 m) antenna in low Earth orbit by the end of this decade. This spacecraft, working with ground telescopes in the USSR, would produce a well-filled aperture but with maximum baselines little longer than can be obtained with current VLBI networks. Thus it is being regarded as a step towards a more ambitious mission which might, like QUASAT, be launched in the 1990s. This second mission resembles QUASAT in several ways: for example, it would employ wider bandwidths and higher observing frequencies. It also offers an appealing way around the trade-off between resolution and image quality which inevitably arise when only one orbiting telescope is used. There would be much to be gained if QUASAT and the Soviet telescope (unofficially dubbed KBACAT) were to be placed in complementary orbits. In the most entertaining few minutes of the entire meeting J.F. Jordan (Jet Propulsion Laboratory) showed a time-lapse movie simulation which demonstrated graphically how, in this regard, the QUASAT-KBACAT baselines would enable us both to have our cake and to eat it, for both longer baselines and a better filled aperture plan are then possible.

QUASAT seems destined to be a competitor in the 1987 selection round in ESA and could get a 'new start' in NASA that same year. The antenna concept is expected to be selected in summer 1985, with a joint NASA-ESA phase A study in 1986. Launch could be in 1992. □

P.N. Wilkinson is at the Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL, UK.

Martin Ryle 1918-1984

SIR Martin Ryle, pioneer of radio astronomy and explorer of the Universe, died on 14 October 1984. His insight and originality in radio science was first displayed in wartime, when he was recruited directly from Oxford in 1939 to work in the Telecommunications Research Establishment. Here his first tasks were to develop airborne radar aerial systems and to test equipment for the new centimetric waveband; these provided an introduction to radio techniques which prepared him well for his vital, but almost unrecorded, contributions to radio countermeasures.

It was Jack Ratcliffe who persuaded Ryle to start radio research in the Cavendish in 1945, with the support of an ICI Fellowship. The wartime discoveries by John Hey, of intense radio emissions from the Sun and from a mysterious source in the constellation of Cygnus, provided challenging subjects, well suited to Ryle's experience in the location of radio transmitters and the recognition of unusual signal characteristics. He devised a radio interferometer, analogous to Michelson's optical counterpart and soon named after it, which showed in 1946 that sunspots emit powerful, circularly polarized radio waves, and that the quiet sun is surrounded by a hot, radio-emitting corona. A succession of students was encouraged to map this coronal emission, using ingenious developments of the interferometer; these developments can now be seen as the first stages of the technique of aperture synthesis, one of Ryle's most famous achievements.

The investigation of the Cygnus radio source followed similar lines. It was a natural hypothesis that this source might be a star that radiated like the Sun but much more powerfully. The successful location of the source, and its identification as a very distant and very unusual galaxy, opened the way for radio astronomy to investigate the structure of the Universe on the largest scale. At that time, there could be no clear interpretation of the many discrete radio sources appearing in the early surveys; they were at first supposed to belong to our Galaxy, where they would account for the background radio emission from the Milky Way. Ryle determined to survey and locate a large number of these sources. For this purpose he built a large interferometer array, working at a wavelength of 3.7 m. The results were astonishing: the sky was apparently crowded with sources, with almost no concentration in the Milky Way.

In his 1955 Halley lecture, Ryle showed that most of these radio sources must be extragalactic and so distant that they gave the first clear evidence which could be used to distinguish between cosmological theories. His results and their interpretation were immediately the subject of controversy. Australian observations did

not agree with the number counts of the radio sources, while the theorists vigorously defended the sinking ship of the Steady State. The conflict could only be resolved through better observations, which Ryle duly provided by improving the interferometer in successive steps, leading eventually to aperture synthesis.

Aperture synthesis, the simulation of a very large radio telescope by the use of multiple interferometers, was Ryle's supreme invention. It was achieved in 1956, when the technical problems of computation were overcome by the use of the pioneering digital computer, EDSAC. Its full development progressed through the 'one-mile' telescope (1964), which used three 18 m paraboloids, giving an angular resolution better than one arc minute, to the well-known 'five kilometer' telescope (1972), working on shorter wavelengths: this telescope gave, for the first time, an angular resolution of one arc second, similar to that of optical telescopes.

The results, both in the form of number counts and as detailed maps of radio galaxies and quasars, firmly established radio astronomy as the main source of direct observational evidence in cosmology. There were, of course, many by-products, most notably Hewish's discovery of pulsars, which led to Ryle and Hewish sharing a Nobel prize in 1974.

Ryle's originality and determination marked the whole development of Cambridge radio astronomy. He was a magnetic figure, with the power of inspiring of a deep loyalty and a social cohesion among a long succession of students. He had a deep social conscience, which led him in his later years to an active and public concern about nuclear power and armaments. He will be remembered with affection by everyone who worked with him.

F. Graham Smith

100 years ago

EARTHQUAKE MEASUREMENT

IN an article on "Earthquakes" in last week's *Nature* (p. 608), Dr. H. J. Johnston-Lavis takes exception to the records of earthquake motion which I have published, on the ground of their complexity, and pronounces the Plain of Yedo unsuitable for earthquake observations.

Now this seems to me to be a very eclectic way of treating earthquakes. We can measure earthquakes only where we find them, and I suppose the first qualification in a site for an earthquake observatory is that there should be plenty of earthquakes. The Plain of Yedo possesses this qualification in a very high degree; and if the disturbances which occur in it are of a very much more complex character than our *a priori* notions about earthquakes may have led us to expect, it is not the Plain of Yedo that is to blame.

University College, Dundee.

J. A. Ewing

From *Nature*, 31, 4; 6 November 1884.

Anthropology

Prehistory of Amazonian Indians

from Robert M. May

TIME was when the popular image of the Amazonian rain forest was of a teeming paradise with great potential for sustained productivity. Following the collapse of many ill-conceived 'development' projects, however, the contemporary image is more one of fragile forests perched on impoverished soils, or even of a Green Hell¹. Consistent with this recent picture is a vision of the Amazonian Indians as simple people eking out a precarious existence, with population densities that have always been too low to sustain any complex social organization or culture. A series of recent studies contest this view of the Amazonian Indian²⁻⁶. Several studies argue that prehistoric population densities were an order of magnitude or more higher than conventional estimates, or than densities found today. Moreover, it is argued that the Indians' knowledge about and use of their environment was much more sophisticated than has usually been believed.

Population

The early estimate of the density of aboriginal Indians in Amazonia before the advent of Europeans was less than one individual per km². As summarized by Dobyns² in an important revisionist article, this estimate is based partly on historic records, partly on assessment of the carrying capacity of the environment and partly on direct observations of contemporary densities of aboriginals. Dobyns discusses each of these three approaches.

First, most historians and anthropologists have tended to discount contemporary sixteenth and seventeenth-century estimates of Indian populations in both North and South America as gross exaggerations. Another revisionist, Myers³, is understanding of such discounting, even though he disagrees with it: "we must recognize that the earliest historical sources were written either by adventurers in search of gold and kingdoms to conquer, or by missionaries whose vast expenditures could only be justified by a proportionate number of conversions to the Catholic faith. When disinterested parties finally did reach the Amazon, they found no evidence of the large populations which were reported to have lived there". Dobyns shows, however, that when the old records — or indirect inferences drawn from them — are examined carefully, and cross-checked one against another where possible, their high estimates of the initial size of aboriginal populations in North America, in Central America, in the Incan Empire and in Amazonia appear to stand up. Discrepancies arise because populations crashed between first contact and the arrival of reliable historians, under the impact of

diseases introduced by Europeans. Likewise, Myers's detailed analysis of all available historical records for the Indians along the Ucayali River in eastern Peru indicates that "large and stable communities were in fact characteristic of the mainstream Ucayali tribes at the beginning of the historic period, but that they collapsed with the precipitous population declines caused by Spanish diseases".

Second, our understanding of the ability of the Amazonian rain forest to sustain humans is not good enough to provide any reliable estimate of population sizes (more on this below).

Third, direct observations of population sizes several years after European contact are not useful, by virtue of the havoc



Buhagana Indian, East Colombia, c.1890. (Peter Newark's Western Americana).

wrought by disease. Such useful observations as do exist come from geographically isolated areas of the Amazon Basin where a few native populations appear to have experienced their first effective contact with Europeans so recently that ethnologists have been able to collect information about the population before contact. One group, first contacted in 1926, was virtually wiped out by influenza in 1929. In another, 300 died of pulmonary oedema in 38 hours. Another group of 300 Sabane came into a Brazilian post, where the wagon bringing gifts and rations for them also brought bronchial pneumonia; the few survivors fled back to their settlements, spreading the epidemic. Anecdotes of this kind make it clear that observations of populations several

decades after first contact give little hint of pristine population densities².

In pursuit of an estimate of the depopulation typically caused by disease, Dobyns pulls together information from various places in the New World and concludes² that over the 100 years following first contact, "the depopulation ratio of 20 to 1 appears to be a sound, if perhaps conservative, tool to employ as a hemispheric minimum". This figure coincides with an earlier and less encyclopedically-based estimate by Borah of a 95 per cent reduction by disease in native populations².

As Myers has emphasized elsewhere, viral and bacterial diseases can travel along trade routes, accompanying items of trade into remote regions long before the first face-to-face contact with Europeans is made⁴. Thus, the first reported contact may be with a native population already severely reduced by disease.

A different and independent indication that Amazonian Indian populations were relatively dense before the arrival of Europeans comes from analysis of sites of 'Indian Black Earth' (*terra preta do indio*)⁵. These patches comprise black earth intermixed with potsherds and stone stools; they are scattered throughout the Amazonian Basin and are classified as a soil type. Suggested explanations of the phenomenon include fallout from volcanoes, sediments from vanished Tertiary lakes, remnants from slash-and-burn agricultural practices and cultural layers accumulated at former Indian villages. Smith argues persuasively that the first three explanations are not consistent with the observed geographical distribution or the depth and composition of the patches⁵. He concludes that the *terra preta* sites are the vestiges of ancient cultures in Amazonia, providing evidence that "precontact native populations were in many cases large and sedentary, particularly along rivers". If this is so, says Smith⁵, "biologists should be wary when they discuss virgin Amazonian ecosystems. Potsherds and black earth may lurk under control plots and pristine nature reserves".

Lifestyle

Not only may the early Amazonian populations have been more dense than previously thought, but it also seems likely that they were correspondingly more sophisticated. In particular, the fieldwork of Posey on the Kayapo Indians in Central Brazil has illuminated these peoples' rich appreciation of their environment⁶. The present population consists of about 2,500 Indians inhabiting 9 villages scattered within a 20,000 km² reserve.

Like most such people, the Kayapo have a large vocabulary to describe the subtleties of their surroundings. Their major subdivisions of the forest include *ba-kamrek* (riverine forest), *ba-epti* (higher forest subject to intermittent flooding every 10 years or so, but containing some richer

alluvial soils), *ba-kati* (*terra firme*, lands that do not flood) and *ba-rarara* (forest with intermittent openings). Kayapo villages are sited to have access to a variety of these ecological zones, and the patterns of field ownership read like something from a current behavioural ecology text: although some fields are selected from *ba-kati*, others are carved from *ba-epti* because the better yields offset the unpredictable losses due to flooding; chiefs' fields and women's collective fields are usually planted in such a way as to minimize total crop loss for any one family group. Not only do the Kayapo utilize some 250 plant species for their fruits, along with several hundred others for their nuts and tubers, but they also practice a transitional form of cultivation. In addition to cultivated plots near the villages, men on hunting treks rely on 'islands of resources' along their vast network of interlinked forest trails; these 'forest fields' are maintained by collecting plants during the day's travels and then replanting them near established campsites. At least 54 species of these semi-domesticated plants are used in the forest fields.

The Amazonian Indians also know their animals: "every significant detail of the life habits of animals is part of [the] hunter's knowledge, including the sound of its cry, its preferred foods, its excrement, its scents, the teeth marks it makes on fruit etc"⁷. Beyond this, the Kayapo practice a transitional form of domestication of animals. Thus, they deliberately stack the remains of banana and palm plants near villages to attract adult beetles that lay eggs in such refuse; in due course, the Indians know when to collect the tasty and nutritious grubs that mature from the eggs. The Kayapo further recognize some 54 species of stingless bees and 2 species of stinging bees: some of these are cultivated for honey and beeswax, and other species are cultivated (in rotting logs in holes dug in fields) as pollinators to increase crop yield.

Although little is known about the medicinal and other properties of the Kayapo's cultivars, at least two merit mention. The red seeds of *urucu* (*Bixa orellana*) are used to flavour foods and also as the main ingredient in body paint. Red body paint is widespread for adornment throughout Amazonia, but Posey also found *urucu* "to be an effective natural insect repellent with a significant reduction (as much as 84 per cent) in insect bites when painted on the body"⁶. The wild ginger *madn-tu* (*Zinziber*) is used — apparently with some effect — as a medicine against intestinal parasites.

Posey takes to task the common misconception that fields cultivated under slash-and-burn agriculture are totally abandoned after a few years. Kayapo 'abandoned' fields continue to produce harvests of yams and taro for 5–6 years, bananas for 12–15 and *urucu* for 20 years or more. The reforestation sequence also produces a variety of foods that attract

game in artificially high densities, and this phenomenon is also exploited. The Kayapo "thus do not have a clearcut demarcation between fields and forest, nor between wild and domesticated. Rather they have a more general system for classification of ecological resources that forms a continuum between wild and domestic ones, all of which figure into integrated management strategies"⁶.

Posey also tells some fascinating tales of the way myth and ritual interweave with natural history facts. One example must suffice. The myth: the little red ants are gentle like women and they are the relative/friend of the manioc; this is why women use the little red ant to mix with *urucu* to paint their faces in the maize festival; the little red ant is the guardian of the field. The facts: manioc has extrafloral nectaries that attract the ants, which in turn trim away any bean vines that would prevent the new and fragile manioc stems from growing; the beans can climb on the maize, which is sturdy enough to shoot up undamaged; and the leguminous beans help

fertilize the total crop. Thus, the coevolutionary relation of ants to the manioc-maize-bean system does indeed facilitate the women's work.

It is tempting to go beyond this, seeing such myth and ritual as functioning to regulate the Indians' use of natural resources. Although this may sometimes be so, I do not always find Posey's argument compelling: the Kayapo 'belief system' may simply serve to describe the world, rather than to manipulate it. Even so, such finely-nuanced description puts these people closer to the contributors to *Nature* than is commonly conceded. □

1. Goodland, R.J.A. & Irwin, H.S. *Amazon Jungle: Green Hell to Red Desert?* (Elsevier, New York, 1976).
2. Dobyns, H.F. *Current Anthropology* 7, 395 (1966).
3. Myers, T. *Ethnohistory* 21, 135 (1974).
4. Myers, T. in *Networks of the Past: Regional Interaction in Archaeology* (eds Francis, P.D., Kense F.J. & Duke, P.G.) 84 (Chacmool Press, Calgary, 1981).
5. Smith, N.J.H. *Ann. Assoc. Am. Geogr.* 70, 553 (1980).
6. Posey, D.A. *Ciencia e Cultura* 35, 877 (1983).
7. Carneiro, R. in *Native South Americans* (ed. Lyman, P. (Little-Brown, Boston, 1974).

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, New Jersey 08544, USA.

Nomenclature

Code for Spanish names of birds

from I.F. Aguillo and M. Fernández-Cruz

ALTHOUGH publication of the International Code of Zoological Nomenclature was one of the most important achievements of modern taxonomy, vernacular nomenclatures remain useful as a communication tool; common names often hold a lot of information, as they represent a particular aspect of the relationship between man and his environment. The use of vernacular names is particularly common in the case of ornithology, a science with a high number of amateurs and much non-scientific literature. In that sense, ornithologists use both common and scientific names in their papers, giving special value to the former as an informative tool.

There is a great vernacular richness associated with birds, in every language and region of the world. The amount of information available is so large and varied, that quite often it is necessary to accept widely used names that are not as informative as they could be. This is not the case with the Spanish-speaking community of ornithologists, probably because the scientific infrastructure of the regions and the ornithological bibliography available in Spanish are very scarce. Thus the situation is favourable for the vernacular nomenclature of birds to be modified so that the names are more meaningful.

Since there are very few lists with accepted names in Iberoamerica, it will first be necessary to make a checklist to be used as a reference by all of the members of the community. Such a project has already been launched but its success will require the international cooperation of all 20 or so

countries integrated in this community. The ultimate aim is to produce a system of common names for the Iberoamerican and European avifaunas, based on associative and meaningful names.

General agreement has led to the proposal for a unified Code of Spanish common names which states the rules for deriving the new nomenclature. Although its criteria are inspired by the International Code of Zoological Nomenclature, informative names will be favoured over priority ones. Names will be brief and easy to write and pronounce; and will be selected on the basis of how informative they are.

The unified nomenclature will be binominal. The first word will be the 'generic name' which refers to all those species that can be grouped in accordance with taxonomical, morphological or historical-usage criteria. Preferably, the name should be derived from the vernacular roots of the community people, adapting it to the Spanish language. In other cases, the name will include information about ecology and ethology of the bird, but not morphological information. The 'specific name' will follow as an adjective, will characterize each species, and represent it in every way. This name will include information about the bird's morphology and appearance, geographical distribution and so on.

The Code and details of the proposed unification project, in Spanish, are available from the authors at: Cátedra de Vertebrados, Facultad de Biología, Ciudad Universitaria, 28040 Madrid, Spain. □

Solid-state perspectives of the photoelectrochemistry of semiconductor–electrolyte junctions

F. Williams* & A. J. Nozik

Solar Energy Research Institute, Golden, Colorado 80401, USA

More than one mechanism is required to account for the photoinduced movement of charge across semiconductor–electrolyte junctions. Distinguishing the various mechanisms, which is here accomplished using a synthesis of theoretical ideas from both solid-state physics and electrochemistry, may have important consequences in research related to potential applications of liquid junction devices in solar energy conversion.

PHOTOEXCITATION of semiconductors in contact with liquid electrolytes, followed by charge transfer of the photogenerated carriers into the liquid, is an important and well-studied phenomenon. This process can occur via several mechanisms which are generally not clearly differentiated in the literature. Furthermore, there are important similarities, as well as differences, between photoexcited semiconductor–electrolyte, semiconductor–vacuum, and semiconductor–solid systems. In this review, we discuss the mechanisms for photoexcited charge injection within a common framework, and connect the phenomena with other well understood semiconductor concepts. In addition, we provide a glossary of terms that joins the languages of solid-state physics and electrochemistry and will help readers of either discipline with the language of the other.

Two major phenomenological regimes can be identified. If the photoexcitation creates charge carriers with energies sufficiently large to be injected into continuum states of the liquid, then this process is defined as photoemission (PE). On the other hand, if the photoexcitation energy is sufficient to bridge the semiconductor band-gap, but insufficient for emission to occur into a continuum state of the liquid, the photogenerated carriers can still be injected into the liquid. This process occurs by charge separation and transport facilitated by the existence of a space charge (depletion) layer at the semiconductor–liquid interface; the process provides the basis for photoelectrochemistry. We call this process photoinduced charge transfer (PCT).

Photoemission

Photoelectric emission from a semiconductor into a vacuum is well understood. The simplest systems for analysis have no surface charge or space charge and their electronic states and transitions are illustrated in Fig. 1. The vertical axis represents total energy of an electron (eigenstate of the electronic hamiltonian) and the horizontal axis is position perpendicular to the surface. Within the semiconductor, the highest occupied band state at equilibrium is the top of the valence band E_v ; the lowest unoccupied state is the conduction band edge E_c ; the two are separated in energy by the forbidden gap E_g . The energy of an electron just outside the semiconductor, with zero kinetic energy, lies above E_c at the photoelectric threshold E_0 , often referred to as the vacuum level. The Fermi level E_f describes the thermodynamic occupancy of all electronic states, including dopant states within E_g .

In the photoemission measurements, the semiconductor is illuminated with photons of energy $h\nu$ and electrons are emitted into the vacuum and collected by a metal at an applied voltage

V with respect to the semiconductor emitter. For the metal there is no band-gap and E_f lies at the top of the occupied band states. For both semiconductor emitter and metal collector the work function W is the energy from E_f to the vacuum level. Two experimental conditions are shown: Fig. 1a for a voltage V_i between collector and emitter just sufficient for all emitted electrons to be collected; Fig. 1b, for a voltage V_e which only collects the most energetic electrons. From these measurements the photoelectric threshold E_0 can be obtained:

$$E_0 = W_e + E_f = h\nu + e(V_e - V_i) \quad (1)$$

where $h\nu$ is the photon energy. The work function of the emitter,

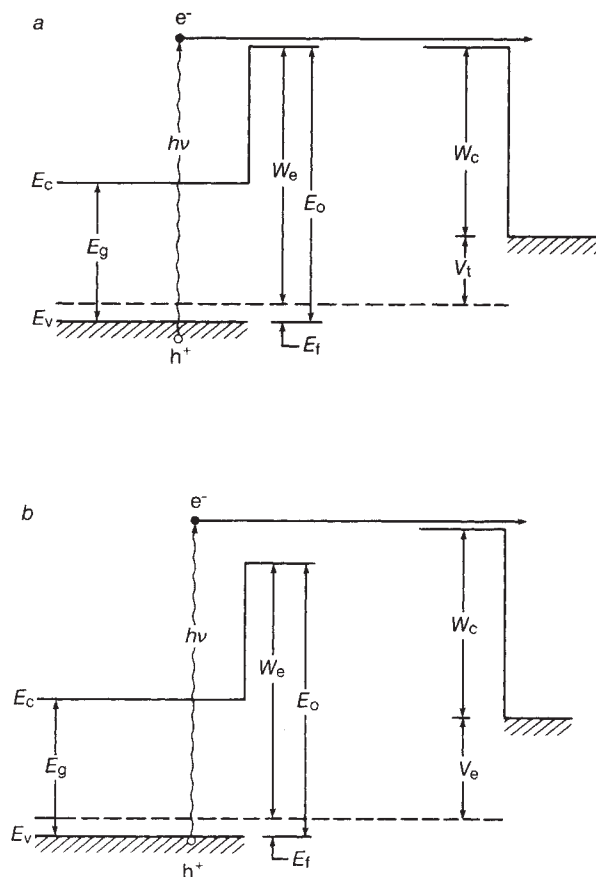


Fig. 1 Energy diagram for photoemission from semiconductors into vacuum: *a*, voltage that collects all emitted electrons; *b*, voltage that collects most energetic electrons.

* Deceased June 1984. Work performed while on sabbatical leave from Physics Department, University of Delaware, Newark, Delaware 19711, USA.

The language of solid-state physics and electrochemistry

Acceptor dopant: An impurity which, in the neutral state, tends to accept an electron from the valence band, that is, to donate a positive hole.

Adiabatic potential: The effective potential for motion of atoms in a molecule or in a condensed phase, obtained in the approximation that the electrons move rapidly and the nuclei slowly. It consists of two terms: the nuclear-nuclear (or core-core) interactions, plus the eigenvalues of the electron state with parametric dependence on atomic coordinates.

Ballistic electron: An electron which traverses the active region of a device without generating or annihilating a lattice vibration (phonon).

Band structure: The relationship between the one-electron eigenvalues and the quasi-momentum which characterizes electronic states for crystals. For semiconductors a gap of forbidden eigenvalues lies between the occupied valence band and the empty conduction band. If the maximum of the valence band and the lowest minima of the conduction band are at the same value of the quasi-momentum, that is, Γ -points, the semiconductor is said to have a direct gap; if not, that is, lowest conduction minima at X- or L-points—an indirect gap.

Band tailing: The density of states for the band structure of undoped, perfectly-ordered semiconductors decreases rapidly to zero at the band edges. Dopants, disorders, space charges and applied electric fields perturb the density of states leading to tailing from the band edges into the forbidden gap.

Depletion region: The space charge region in a semiconductor which is depleted of majority carriers and produces a potential barrier for their transport to the interface.

Donor dopant: An impurity which in the neutral state tends to donate an electron to the conduction band.

Effective mass: Electronic particles in crystals react to an applied force as if they have an inertial mass which is different from the free-electron inertial mass. This mass is designated as the effective mass. The difference between the effective mass and the free-electron mass is due to the interaction between the electronic particle and the periodic potential of the crystal.

Electron-phonon interaction: The separation of the many-body problem of electrons plus nuclei in crystals into electronic band structure on the one hand and lattice dynamics (phonons) on the other is an approximation. A correction to this approximation involves coupling between the electrons in their states with the lattice dynamics and is called electron-phonon interaction. This interaction is essential for thermalization between electronic particles near their band

edges and for the molecular or lattice dynamics. It is also responsible for the polaron.

Exciton: An excited electronic state of crystals in which the excited electron and the positive hole are coupled (electron-hole pair); no net charge is transported during exciton motion. The effective mass exciton is a coupled conduction band electron and valence band hole: thus optical creation of effective mass excitons occurs on the long wavelength edge of the band-to-band optical absorption.

Fermi level: Energy level for which the occupational probability for an electron is 0.5, under thermal equilibrium conditions. The occupational probability for a positive hole is also 0.5 at equilibrium.

Franck-Condon principle: The most probable optical transitions between electronic states occur on a time scale short compared to the period of atomic motion, that is, molecular or lattice vibrations: the positions and momenta of the atoms remain fixed during these transitions. Franck-Condon transitions occur vertically between adiabatic potential curves for the initial and final electronic states of the transitions.

Helmholtz layer: The double layer of charge created at a solid-liquid electrolyte interface by the adsorption on the electrode of ionic and/or molecular species from the liquid.

Hot carriers: Electrons or holes that are not in thermal equilibrium with the lattice and have energies above the respective band edges, in the form of excess kinetic energy.

Negative electron affinity: Effect produced when sufficient band bending is generated in a p-type semiconductor with a low work function surface layer such that the position of the conduction band edge in the semiconductor bulk is above the vacuum level. Thus, photoemission of electrons into vacuum by band-gap photons can occur if this happens faster than intraband relaxation in the depletion region.

Polaron: For polar semiconductors, conduction band electrons and valence band holes do not move at their respective band edges as pure electronic particles but rather carry with them a cloud of lattice polarization (virtual phonons). The electronic particle bearing lattice polarization is itself particle-like and is called a polaron. The polaron mass is in general greater than the band effective mass, and quantitatively dependent on the electron-phonon interactions.

Positive hole: The ground state of a pure semiconductor has a filled valence band and an empty conduction band. States of excitation are thus best described focusing attention on the small concentration of electrons in the conduction band and/or the small concentration of unoccupied valence band states. The latter are electronic particles called positive holes, often abbreviated to 'holes'.

W_c is obtained from the first of these measurements if the work function of the collector, W_c , is known¹.

The effects of space charge are exemplified by negative electron affinity photocathodes², shown for p-GaP in Fig. 2. Equilibration with several monolayers of n^+ -Cs₂O yields a large degree of band bending in the p-GaP, arising from the difference in work functions of these two materials and yielding a space charge. The eigenstates of the complete electronic hamiltonian, including the field due to space charge, are now more complicated than those described earlier, so that the position-dependent band edges E_c and E_v in the space charge region have a somewhat different interpretation than E_c and E_v in the space-charge-free region. In both regions, conduction electrons and positive holes (empty states in the valence band) which are in thermal equilibrium with the lattice move at energies near E_c and E_v , respectively. Band-to-band photoexcitation with photons bigger than E_g produces electrons with energies above the conduction band edge in the bulk. These electrons relax rapidly in <10 ps, in the flat-band region, and diffuse as thermalized carriers to the depletion region. From within this region, the photogenerated carriers can be emitted, without relaxing, to the conduction band edge E_c at the surface. Thus, electrons with energies in excess of the energies of thermal carriers in the depletion region, that is, non-equilibrated with the lattice at the GaP/Cs₂O interface, are emitted into the vacuum. Such electrons are one type of non-thermalized electrons, generally referred to as 'hot' electrons. Their ability to cross the depletion layer without thermalizing in that region, in contrast to the situation

in the flat-band region, results from the existence of a narrow depletion layer (<200 Å) and the associated large electric field.

Photoemission into a vacuum is irreversible because of the infinity of states in free space. Photoemission from a metal into a vacuum can be obtained from the above analysis by taking the limit $E_g \rightarrow 0$, where E_g is the band-gap.

Photoemission from a semiconductor into another phase of condensed matter is a valid concept if the electron emission is into continuum states, that is, into the conduction band, for example, of an aqueous electrolyte. The efficiency of emission is limited by differences in effective momentum k and mass m^* of the two phases. The concept can be generalized to include positive hole photoemission from the semiconductor into a continuum of hole states. These processes are illustrated in Fig. 3 for photons labelled $h\nu_{PE}$. The band edges and energy gap for the aqueous electrolyte are based on the model for liquid water as a large gap, lone-pair, amorphous semiconductor³. Figure 3 shows the effects of a space charge arising from the inequality of the work functions of semiconductor and electrolyte, $W_{sc} \neq W_{el}$. For the p-type semiconductor, $W_{el} < W_{sc}^p$; for the n-type semiconductor, $W_{el} > W_{sc}^n$. These are the usual cases since $W_{sc}^p \approx W_{sc}^n + E_g$.

Electron or hole emission involves interchange between kinetic and potential energy; only total energy remains constant during the emission processes shown in Figs 1-3. The total energy for each electron divides as follows: the energy from the Fermi level, E_f , up to E_c is potential energy; from E_c to the eigenvalue E it is kinetic energy. Thermodynamically, the energy

Quantum well: An effective potential well created by a minimum in the conduction band or a maximum in the valence band that arises when a smaller band-gap semiconductor is sandwiched between a larger band-gap semiconductor. When the physical dimensions of the sandwich layers become small (comparable to the mean free path of electrons or holes), conduction electrons and positive valence band holes are localized in a dimension perpendicular to the layers because of quantization of the energy states in the well.

Quasi-Fermi level: The energy level for which the occupational probability for an electronic particle is one-half under steady-state conditions (time-independent but non-equilibrium). The quasi-Fermi levels for electrons and positive holes are different because of non-equilibrium concentrations of injected electronic particles, either electrons and/or holes.

Redox couple/redox potential: A redox couple is a pair of electron donor and acceptor species in solution that define a potential for charge transfer. At equal concentrations of donor and acceptor species (called the reduced and oxidized forms of the couple), the thermodynamic driving force for electron transfer to the acceptor or from the donor are equal. The couple is in equilibrium and this condition defines the standard redox potential, E_{redox}^0 , which is the electrochemical potential in the liquid electrolyte. The potential at other concentrations is related to E_{redox}^0 through the Nernst equation.

Solvated electron: A conduction electron in an electrolyte polarizes nearby molecules so that in this region the structure of the electrolyte is deformed and the electron acquires close ligands, that is, it is solvated. The solvated electron is, in fact, a polaron of rather large mass.

Space charge: The double layer of charge produced near the surface of a semiconductor in contact with a second phase resulting from the initial inequality of electrochemical potentials (that is, work function, redox potential, Fermi level) for the two phases. The charge is generally distributed over finite dimensions in the semiconductor.

Thermalized carriers: Electrons or positive holes that are in thermal equilibrium with the lattice, and move in semiconductors near the band edges E_c and E_v .

Work function: Work required to remove an electron from the Fermi level of a condensed phase into vacuum with the electron having zero kinetic energy. For liquid electrolytes the work function is identified with the redox potential of the electrolyte.

Zero-phonon transition: Optical transition between electronic states in which phonons are neither created nor annihilated. The transition normally occurs from the zero-point vibrational level of the initial state to that of the final state.

available to do useful work depends on an electron occupancy of state E in excess of thermal occupancies at the lattice temperature, and on $E - E_f$; the total available energy is independent of division into kinetic and potential energy. Similar considerations apply to the positive holes. With applied voltage and/or photoexcitation, E_f may for some conditions be described as split into two quasi-Fermi levels for electron and positive hole, E_f^e and E_f^h , respectively. The concept of quasi-Fermi levels is valid if occupancies of electron and hole states are separately in accordance with Fermi statistics.

Photoemission from the semiconductor into the forbidden gap, E_g^{el} , of the electrolyte modelled as an intrinsic semiconductor (undoped and free of defects) is obviated by the absence of allowed states within E_g^{el} . Here the electronic wavefunction ϕ attenuates exponentially in the electrolyte with a mean distance $\hbar/\sqrt{2m|E'|}$, where E' is measured from the band edge. Thus photoemission with electron energies within E_g^{el} can occur only within a range of a few molecular distances.

A more complete model of an aqueous electrolyte takes account of band tailing: that is, the allowed eigenstates tail into the forbidden gap. This shows itself in the exponential tail of the optical absorption of amorphous semiconductors and is evident in the vacuum ultraviolet absorption of water³. In general, band tailing can be explained in terms of the microfields generated by departures from a perfect periodic potential so that sharp E_c and E_v no longer exist⁴, or, alternatively, in terms of band edges which undulate spatially because of compositional and/or structural inhomogeneities⁵. Thus photoemission into

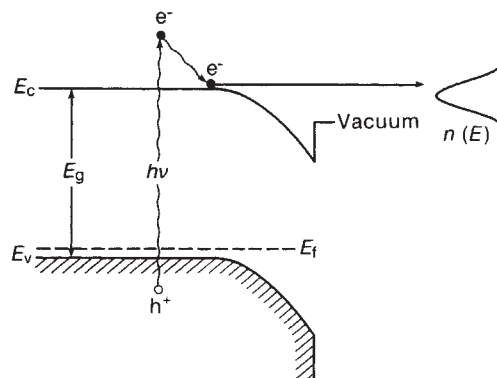


Fig. 2 Hot electron emission into vacuum resulting from band bending in negative electron affinity photocathode. Energy distribution of emitted electrons is shown (from ref. 2).

the band tails is predicted but with lower probability because of the low densities of states.

Another mechanism specific to liquid electrolytes is photoemission directly into solvated electron states⁶. This depends on the molecular polarization and orientation in the vicinity of the trapped electron. The lowest of these states in water lies about 2 eV below the unsolvated electron states. The solvated electron is a low mobility polaron, a quasi-particle well known in the solid state. Because electron emission is fast compared to intermolecular motion, photoemission directly into solvated electron states depends on the electrolyte preparing itself spontaneously in a state of polarization consistent with the solvated state. Since liquid electrolytes are at equilibrium, the probability of such pre-transfer polarization into states 2 eV below the band edge is quite unlikely. Similar considerations apply to the possibilities for photoemission into self-trapped positive hole states of the electrolyte.

Photoinduced charge transfer

Photoinduced charge transfer (PCT) is the process wherein electron-hole pairs are first created in the conduction and valence bands of semiconductor electrodes by absorption of photons with band-gap energies; then separated by the field in a depletion layer at the semiconductor-electrolyte interface; finally electrons or holes are injected into the electrolyte to drive redox reactions. For n-type semiconductors, the holes move to the semiconductor-electrolyte interface to generate an oxidation reaction (semiconductor is a photoanode); for p-type semiconductors, electrons move there to generate a reduction reaction (semiconductor is a photocathode). Photoinduced electron transfer is illustrated in Fig. 3a and hole transfer in Fig. 3b for photons labeled $h\nu_{\text{PCT}}$. PCT can occur from carriers with energies near the band edges in contrast to the energy required for photoemission. In addition, under certain conditions, fast PCT by direct injection into dopant or defect states of the electrolyte may be possible. Processes involving PCT have been extensively studied in recent years, and many reviews are available⁷⁻¹⁰.

The impetus for this research is the potential application of liquid junction devices to solar energy conversion⁷⁻¹⁰. Photoelectrochemical cells containing photoactive semiconductor electrodes can be used either to drive redox reactions (for example, water splitting, CO_2 reduction, and N_2 reduction) producing fuels and chemicals, or in electrochemical photovoltaic cells to produce electricity.

The important distinction between PCT and photoemission is that in the latter the photogenerated electrons or holes are injected into the continuum states of the electrolyte, whereas in PCT the photogenerated carriers are injected into dopant or defect states that exist in the band-gap of the electrolyte. Dopant states are associated with the redox potential of redox couples added to the liquid, whereas defect states are redox potentials associated with the liquid itself. The latter are the native defects of the liquid electrolyte, for example, H^+ and OH^- ions in liquid

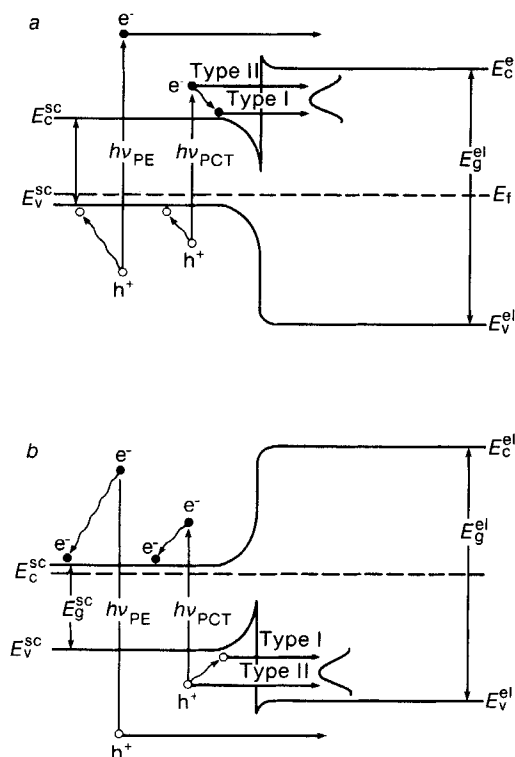


Fig. 3 Photoemission and photoinduced charge transfer (PCT) processes from semiconductors into liquid electrolytes: *a*, for electrons; *b*, for holes. Photoemission occurs with $h\nu_{PE}$; PCT occurs with $h\nu_{PCT}$. The PCT processes are shown for Type I and Type II hot carrier injection. Thermalized PCT processes inject the carriers from the bottom of the semiconductor conduction (or valence) band at the semiconductor-liquid interface. The energy distribution of the electron acceptor in the electrolyte is shown, and arises from fluctuations of the solvation shell around the acceptor (see Fig. 4 for further details).

water. Thus, the distinction is based primarily on the nature of the relevant electronic states of the electrolyte, rather than on intrinsic differences in the nature of the photoprocesses in the semiconductor. This is evident in Fig. 3.

In PCT, the carriers injected into the electrolyte can either be fully thermalized in the semiconductor or not. When thermalized, the photogenerated electrons and holes arrive at the interface with energies corresponding to the band edges E_c and E_v at the surface, respectively. These carriers always remain in equilibrium with the lattice through electron-phonon interactions, and dissipate the kinetic energy created by band-bending in the depletion layer as heat, as the carriers traverse the depletion layer. In thermalized PCT, only redox potentials that lie within E_g of the semiconductor at the interface are accessible to the injected charge.

For the case of non-thermalized PCT, illustrated in Fig. 3, the photogenerated carriers are injected into the electrolyte before they have undergone complete thermalization as they move from bulk semiconductor into the electrolyte. This process, called hot carrier injection, has been analysed in some detail¹¹⁻¹⁴, and is discussed further below. For hot carrier injection, redox couples that have potentials outside E_g of the semiconductor at the liquid contact are accessible.

Finally, we discuss the role of surface states. In semiconductor-vacuum and semiconductor-solid interfaces, the major effect of surface states is to pin the Fermi level so that the band-bending is independent of, or insensitive to, the nature of the second phase contacting the semiconductor. In semiconductor-liquid interfaces, the nature and effects of surface states can be quantitatively and qualitatively different from those of solid state junctions. First, ionic species in the electrolyte can be specifically

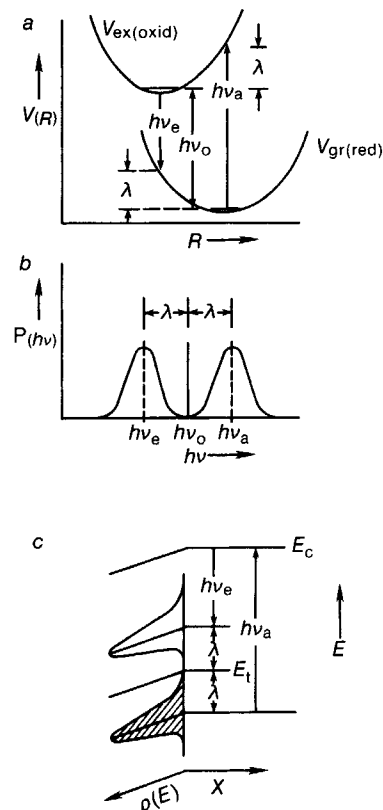


Fig. 4 *a*, Adiabatic potentials for ground (reduced) and ionized (oxidized) states of dopants in polar semiconductors, insulators or electrolytes, with Franck-Condon and zero-phonon transitions. *b*, Probability distributions for radiative transitions. *c*, Electronic energy levels, with thermodynamic dopant level E_t and effective densities for transitions. For redox couples in electrolytes E_t is identified with E_{redox}^0 (see box).

adsorbed onto the semiconductor either to passivate surface states or create new ones¹⁵. Second, Fermi level pinning or surface state charging under illumination can move the relative energy levels of the semiconductor with respect to those of the liquid electrolyte, and hence affect the kinetics of charge transfer¹⁶⁻¹⁸. Finally, surface states can mediate the charge transfer such that it occurs from the surface state rather than from the semiconductor band edges²⁰; this could also have a very large effect on the kinetics of PCT.

Large-gap amorphous semiconductors

Liquid electrolytes such as water can be described, with respect to electronic states, as large-gap amorphous semiconductors. The optical properties of pure and doped liquid water have been explained with this model³. It also provides the basis for interrelating some electrochemical concepts (for example, redox couples) with semiconductor concepts (for example, donor and acceptor dopants). However, semiconductor theoretical analyses of electrolyte-semiconductor junctions require modification to take account of the large density of mobile ions in liquid electrolytes.

Charged point defects in crystalline semiconductors or insulators are well known to have bound states for electronic particles: bound electron states for a positively-charged defect and bound hole states for a negatively-charged defect—termed donor and acceptor dopants,²¹ respectively. Some uncharged dopants bind an electron or a positive hole and are referred to as isoelectronic dopants²².

These dopants in polar semiconductors or insulators have bound states whose energies are dependent on occupancy. This is a consequence of electron-phonon interaction and can be

visualized in detail by considering the adiabatic potentials $V(R)$ for two states of occupancy, as shown in Fig. 4a. $V(R)$ is the interaction potential between the atoms, in the region of the dopant with position coordinates $R_1, R_2, R_3, \dots$, symbolically represented by the single coordinate R , and is dependent on electronic state. Thus, the equilibrium value for R is different depending on whether or not an electronic particle is bound to the dopant. Optical attachment or detachment of the particle in accordance with the Franck-Condon principle (see standard texts, for example, ref. 23, and box) will be followed by lattice relaxation from the equilibrium position of the initial state to that of the final state.

The $V(R)$ are assumed harmonic with the same force constant for the ground and excited electronic states; the change in polarization energy is designated λ (anticipating identification with change in solvation energy for ions in electrolytes); $h\nu_0$ is the zero-phonon transition energy and equals E_i , which determines the thermodynamic occupancy of the two states; $h\nu_a$ and $h\nu_c$ are the Franck-Condon transitions for optical absorption (ionization) and emission (deionization), respectively. In Fig. 4b the probability distribution for these transitions is shown. Their band widths and Stokes shift are consequences of zero-point and thermal occupancy of the initial state, and of the transformation from the R -coordinate to transition energy.

These states and transitions can be approximately represented on an electronic energy level diagram as shown in Fig. 4a. We emphasize, as have others²⁴, the differences between $V(R)$ and electronic energy E . E_i lies at that energy below the conduction band edge E_c , with the approximation that E_c is independent of local polarization. For Franck-Condon transitions, the effective densities of states, $\rho(E)$, for ionization and de-ionization are respectively displaced, deeper and shallower by λ , and also broadened. The $\sqrt{E - E_c}$ dependence of $\rho(E)$ for the conduction band is also shown.

The same phenomena occur with redox couples in liquid electrolytes. The two states of occupancy correspond to reduced and oxidized forms of the couple, that is, dopant, similar to those shown in Fig. 4. The structure of the liquid does not have translational periodicity and therefore its lattice dynamics are not describable by phonons; however, the relaxation process is describable in terms of local modes and is basically the same as if the matrix for the dopant were crystalline. The $V(R)$ for each electronic state of occupancy or of excitation is characterized by different equilibrium molecular coordinates R , now in the solvation shell; λ is the energy associated with the reorganization of the solvent molecules around the redox species as it becomes oxidized or reduced. In general, the force constants are different but here are approximated as the same. The energy distributions of oxidized and reduced forms of a redox couple can be related respectively to those of ionized donor (or neutral acceptor) and those of neutral donor (or ionized acceptor) of point charged semiconductor dopants. However, the non-radiative charge transfer transitions for redox couples in electrolytes have different energy dependences than radiative charge transfer transitions for dopants in semiconductors if the former involve prior reorganization of the coordination shell: that is, if the transfer occurs adiabatically rather than vertically. This is, of course, basic to the Marcus theory²⁵. On the other hand, Franck-Condon effects have been effectively invoked by Gerischer²⁶ to describe charge transfer at metal-electrolyte and semiconductor-electrolyte junctions. Tunnelling of electrons or positive holes from electrolyte to semiconductor or metal in accordance with the Franck-Condon principle, that is with fixed atomic coordinates, will be at the energies and with the band widths of Fig. 4.

The transfer of electronic charge between a redox couple and an electrode involves only the states of the couple, not those of either free or solvated electrons, and thus the thermodynamic potential of the couple, E_i of Fig. 4c, will be at the Fermi level of the electrode. Ionic transport maintains continuity of current to the redox couple. This differs from charge transfer in semiconductors, for which Fermi levels are determined by conduction

and valence band edges, as well as by dopant states; these band states are necessary for current continuity because of the absence of significant ionic transport in most semiconductors. The explanation of this difference is that the dopants in liquid electrolytes are fully compensated by other dopants or by native defects. Thus conduction electrons or valence band holes are not present in normal electrochemical processes; they are unnecessary in view of the presence of mobile ions. In other words, for PCT the redox couple is decoupled from the band structure of the electrolyte, whereas the semiconductor dopant is not decoupled from the band structure of the semiconductor. However, for photoemission the relationship of the states of the redox couple to the band structure of the electrolyte is, of course, important; in this case the redox couple is not decoupled from the electrolyte band structure.

The difference between semiconductor-electrolyte junctions and semiconductor-semiconductor junctions is thus evident for PCT. For semiconductor-electrolyte junctions, electrons or positive holes can be injected into redox couples and continuity of current satisfied by ionic transport; for semiconductor-semiconductor junctions electron or positive hole injection from conduction or valence band states of one semiconductor into dopant states within the gap of the other semiconductor is obviated by the absence of continuity of current owing to the immobility of the dopants and the absence of inter-dopant charge transport. However, for photoemission the relationship of the states of the redox couple to the band structure of the electrolyte is, of course, important.

Finally, semiconductor-liquid electrolyte junctions involve space charge, in the electrolyte from accumulation or depletion of mobile ions, and in the semiconductor from electronic charge transport for equilibration¹⁴. The electrolyte is charge compensated so that its Fermi level lies near the middle of its energy gap. For moderate electrolyte concentrations, the mobile ions lead to a narrow space charge, that is, Helmholtz layer, similar, as regards high density of space charge, to a junction involving a semiconductor doped to degeneracy. Dilute electrolytes lead to a wider space charge analogous to low doped semiconductors.

Hot carrier photoinduced charge transfer

Mobile electronic particles are normally equilibrated with the lattice modes; thus, occupancy of the electronic states is characterized by the same temperature as that for occupancy of phonon levels. Exceptions occur under severe perturbations, for example, high electric fields, during which energy is acquired by the electronic particles faster than thermalization with the lattice occurs. In this way the energies of the electrons and/or holes increase and the particles are referred to as hot although their energy distribution may not in general be describable by a temperature. If the density of the carriers is high, that is, $\approx 10^{14} \text{ cm}^{-3}$, collisions among carriers are sufficiently frequent for the distribution to become maxwellian²⁷; otherwise the carrier distribution is non-maxwellian. A well-known example is the displaced maxwellian for an electron gas in an applied electric field, in which the normal velocity distribution is shifted for its velocity component in the direction of the field. Photogenerated hot carriers have, of course, initial distributions dependent on the spectrum of the photoexcitation: for example, with a monochromatic source and excitation limited to transitions with $\Delta k = 0$, the initial photo-electron and photo-hole distributions will each be a δ -function. For dilute electrons in intense electric fields, over distances less than the mean free path for phonon scattering, ballistic motion occurs. This phenomenon has been reported for short semiconductor devices²⁸.

The energy distributions of photogenerated conduction electrons and valence band holes, in the region of a semiconductor-electrolyte junction, depend on the photoexcitation spectrum, on intraband relaxation, on interband electron-hole recombination, on the rate of charge transfer from semiconductor to electrolyte, and on the dependence of that rate on electronic energy. The magnitudes of these rates have been estimated, with attention to the effects of quantization of electronic states in

narrow depletion layers on intraband relaxation^{11,13}, and hot carrier transfer predicted from states below the band edge of the bulk semiconductor but higher than thermal energy above the edge at the interface. These are defined as Type I hot carriers. Type II hot carriers are distinguished as having energies more than thermal energy above the band edge of the bulk semiconductor. These processes are illustrated in Fig. 3.

Transfer of photoexcited carriers across semiconductor-electrolyte interfaces have been analysed using the master equation¹³. The total current depends on charge transfer versus recombination. The energy distribution of the charge transferred depends, in addition, on relaxation, density of depletion layer states, and the energy 'window' through which transfer occurs to the redox couple. Back transfer is assumed negligible in this analysis.

Experimental evidence for Type I hot electron injection across illuminated semiconductor-electrolyte interfaces has been obtained for p-GaP (ref. 29) and p-InP (ref. 30) electrodes in non-aqueous electrolyte. In these experiments, reduction reactions of electron acceptor species having redox potentials above E_c at the surface are observed. Such reactions are called supra-band-edge redox reactions, and can only occur via a hot carrier process if the semiconductor band edges do not move with respect to the electrolyte redox levels, during the experiment. The absence of band-edge-movement was established through monitoring of the band-edge energies by capacitance measurements. For p-GaP, the electron acceptor was anthracene (0.9 eV above E_c), and for p-InP, the electron acceptor was p-nitrobenzonitrile (0.5 eV above E_c).

Quantization in small particles

Size quantization effects, previously discussed for narrow depletion layers in planar semiconductor-electrolyte junctions^{11,13}, and well established for semiconductor quantum wells^{31,32} are also evident for small semiconductor particles (20–200 Å diameter) suspended in electrolytes^{33–38}. For the latter, the conduction and valence bands are separately fully quantized in three dimensions into discrete states, and the effective band gap increased as the particle size becomes very small. This essentially produces a hot carrier effect in that the kinetic energy is increased by the confinement. The effect is dependent on effective mass, and is thus material dependent; size quantization effects appear with larger particle sizes for semiconductors with smaller electron effective mass. For the same material, the confinement effect is different for different extrema of the band structure³⁸; for example, the low effective mass Γ -conduction band minimum of tetrahedral semiconductors is perturbed more than the higher effective mass X- and L-band minima. For quite small particles of direct-gap tetrahedral semiconductors, size quantization perturbs the effective band edges into the non-parabolic region of the Γ minimum. For very small particles, for example, ~20 Å for CdS, degeneracy of the three conduction band minima is predicted³⁸.

In addition to the separate perturbations of the conduction and valence bands, small dimensions may result in correlation between the electron and positive hole when they are confined to the single particle. One consequence is that the internal energy of excitons created during photoexcitation increases as the confinement dimensions (D) become comparable to or less than the Bohr radius of bulk excitons (r_B), for example $D \leq r_B \approx 100$ Å for GaAs (refs 36–38). This effect is well established for laminar semiconductor quantum wells³⁹. For small semiconductor particles suspended in electrolytes, the effect of correlation on the band structure itself also becomes appreciable as a result of the electron and hole being confined to a single particle. The effects of correlation on band structure and excitons both depend on occupancy of states, and are present or not depending on the details of the optical experiment; for photoexcitation, the accessible electronic states, whether originating from band or excitonic states, are affected by the Coulomb correlation of electron and hole; for photoemission or fast PCT, the accessible

states are, to a good approximation, band states quantized by confinement, but quite distinct from excitonic states.

In general, for semiconductors with low effective mass electrons, the confinement perturbation is greater than the Coulomb correlation. Thus, for photoexcitation, photoemission or PCT the general features of bulk band structure are predicted to survive in quite small particles, with the forbidden gap increased.

Some of these and related effects have already been found experimentally, for example, the predicted blue shift in absorption spectra^{33,36,37}. However, the distribution of particle sizes in currently available experimental samples gives rise to broad spectra, rather than the discrete spectra predicted theoretically for particles of uniform size and shape.

Size quantization effects with small semiconductor particles in electrolytes should lead to important modifications in the nature of charge transfer and the attainable power conversion efficiencies in photochemical devices: for example, shifts in effective photochemical redox potential have been predicted^{33–38}. Also, unusual kinetics are expected with photoexcitation of very small particles: for example, electron-hole recombination will occur in pairs (geminate recombination), with each pair confined to a particular particle, rather than by the second-order process of bulk semiconductors³⁸.

Conclusions

Photoelectron emission and photoinduced charge transfer at semiconductor-electrolyte junctions can be clearly distinguished. These junctions can be described by semiconductor heterojunction theory if modifications are made to include the effects of charge compensation and of the mobile ions in the electrolyte. Dopants in polar semiconductors and redox couples in electrolytes are similar as regards polarization changes on occupancy but differ as regards equilibration with band edges. Surface states can play an important role in semiconductor-electrolyte junctions in that in addition to the possibility of pinning the Fermi level as in solid state junctions, they may interact strongly with the electrolyte, cause band movement with respect to the electrolyte redox potential, and mediate charge transfer to the electrolyte. Hot electrons were predicted theoretically and found experimentally. Two types of hot carriers can be distinguished on the basis of their electronic kinetic energies in depletion layers and in the semiconductor bulk, respectively. Quantization effects can occur in narrow depletion layers and in small semiconductor particles. This can cause major changes in optical properties and charge transport.

This work was funded by the US Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, under Contract EG-77-C-01-4042. We acknowledge helpful comments by B. A. Parkinson and J. A. Turner. This paper is dedicated to the memory of Ferd Williams.

- Eden, R. C., Moll, J. L. & Spicer, W. E. *Phys. Rev. Lett.* **18**, 597–599 (1967).
- Williams, B. F. & Simon, R. E. *Phys. Rev. Lett.* **18**, 485–487 (1967).
- Williams, F., Varma, S. P. & Hillenius, S. J. *chem. Phys.* **64**, 1549–1554 (1976).
- Redfield, D. *Phys. Rev.* **140**, A2056–2058 (1965).
- Inglis, G. B. & Williams, F. *Phys. Rev. Lett.* **25**, 1275–1278 (1970); *J. non-cryst. Solids*, **5**, 313–327 (1971); *Phys. Rev. B* **11**, 2200–2204 (1975).
- Krohn, C. E. & Thompson, J. C. *Phys. Rev. B* **20**, 4365–4367 (1979).
- Bard, A. J. *Science* **207**, 139–144 (1980).
- Nozik, A. J. *A. Rev. Phys. Chem.* **29**, 189–222 (1978).
- Wrighton, M. S. *Acc. chem. Res.* **12**, 303–310 (1979).
- Heller, A. *Acc. chem. Res.* **14**, 154–162 (1981).
- Boudreaux, D. S., Williams, F. & Nozik, A. J. *J. appl. Phys.* **51**, 2158–2163 (1980).
- Nozik, A. J., Chance, R. R., Boudreaux, D. S. & Williams, F. *Adv. Chem. Ser.* **184**, 155–171 (1980).
- Buhks, E. & Williams, F. in *Photoelectrochemistry: Fundamental Properties and Measurement Techniques*, *Electrochem. Soc. Symp. Proc.* (eds Wallace, W., Nozik, A. J., Deb, S. K. & Wilson, R.) 1–13 (Electrochemical Society, 1981).
- Williams, F. & Nozik, A. J. *Nature* **271**, 137–139 (1978).
- Parkinson, B. A., Heller, A. & Miller, B. *Appl. Phys. Lett.* **33**, 521–523 (1978).
- Bard, A. J., Bocarsly, A. B., Fan, F. R. F., Walton, E. G. & Wrighton, M. S. *J. Am. chem. Soc.* **102**, 3671–3677 (1980).
- Kelly, J. J. & Memming, R. J. *electrochem. Soc.* **129**, 730–738 (1982).
- Cooper, G., Turner, J. A., Parkinson, B. A. & Nozik, A. J. *J. appl. Phys.* **54**, 6463–6473 (1983).
- Kautek, W. & Gerischer, H. *Ber. Bunsenges. phys. Chem.* **84**, 645–653 (1980).
- Chazalviel, J. N. *J. electrochem. Soc.* **129**, 963–969 (1982).
- Kittel, C. *Introduction to Solid State Physics*, 5th edn, 231–237 (Wiley, New York, 1976).
- Thomas, D. G., Hopfield, J. J. & Frosch, C. J. *Phys. Rev. Lett.* **15**, 857–860 (1965).
- Williams, F., Berry, D. E. & Bernard, J. E. *Radiationless Processes* (ed. DiBartolo, B.) 1–38 & 409–430 (Plenum, New York, 1980).

24. Albery, W. J. *A. Rev. phys. Chem.* **31**, 227–263 (1980).
25. Marcus, R. A. *Faraday Disc. Chem. Soc.* **74**, 7–15 (1982).
26. Gerischer, H. *Z. phys. Chem.* **26**, 223–247 (1960).
27. Fröhlich, H. & Paranjape, B. V. *Proc. phys. Soc.* **B69**, 21–32 (1956).
28. Eastman, L. F. *Festkörperprobleme* **22**, 173–187 (1982).
29. Turner, J. A. & Nozik, A. J. *Appl. Phys. Lett.* **41**, 101–103 (1982).
30. Cooper, G., Turner, J. A., Parkinson, B. A. & Nozik, A. J. *J. appl. Phys.* **54**, 6463–6473 (1983).
31. Dingle, R. *Festkörperprobleme* **15**, 21–48 (1975).
32. Dingle, R., Wiegmann, W. & Henry, C. H. *Phys. Rev. Lett.* **33**, 827–830 (1974).
33. Rossetti, R., Nakahara, S. & Brus, L. E. *J. chem. Phys.* **79**, 1086–1088 (1983).
34. Brus, L. E. *J. chem. Phys.* **79**, 5566–5571 (1983).
35. Rossetti, R., Ellison, J. L., Gibson, J. M. & Brus, L. E. *J. chem. Phys.* (in the press).
36. Brus, L. E. *J. chem. Phys.* (in the press).
37. Micic, O. I., Nenadovic, M. T., Nozik, A. J., Rajh, T. & Williams, F. *J. phys. chem.* (submitted).
38. Williams, F. & Nozik, A. J. (in preparation).
39. Bastard, G., Mendex, E. E., Chang, L. I. & Esaki, L. *Phys. Rev.* **B26**, 1974–1979 (1982).

ARTICLES

Detection of a radio emission at 3 kHz in the outer heliosphere

W. S. Kurth*, D. A. Gurnett*, F. L. Scarf† & R. L. Poynter‡

* Department of Physics and Astronomy, The University of Iowa, Iowa City, Iowa 52242, USA

† TRW Space and Technology Group, One Space Park, Redondo Beach, California 90278, USA

‡ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

A radio source in the outer heliosphere has been detected by the plasma wave receivers on Voyagers 1 and 2. The radio emission is observed in the frequency range 2–3 kHz, and is above the local solar wind electron plasma frequency whenever supporting plasma density data are available. The maximum spectral density of the emission recorded is $\sim 10^{-14} \text{ V}^2 \text{ m}^{-2} \text{ Hz}^{-1}$. The bandwidth of the radio noise is $\sim 1 \text{ kHz}$. Possible sources include continuum radiation from Jupiter's distant magnetotail and radiation at the second harmonic of the plasma frequency at the heliopause. If the latter interpretation is correct, these data represent the first remote observations of the heliopause.

THE Voyager 1 and 2 spacecraft are currently sampling the interplanetary environment at heliocentric radial distances of 19 and 13 AU, respectively. Since the Saturn encounters in 1980 and 1981, the plasma wave data have shown very little activity in the outer heliosphere. The primary waves observed include occasional electron plasma oscillations and ion-acoustic waves, sometimes detected in association with interplanetary shocks. Figure 1b shows the trajectories of Jupiter, Saturn, Uranus and Voyagers 1 and 2 for the interval 30 August 1983 to 21 February 1984; during this time a radio emission near 3 kHz was being detected by the plasma wave instruments on both Voyagers. Figure 1a shows a meridional view of the outer planets and the trajectories of the two Voyagers. In this view, the abscissa is in the ecliptic plane and the positions of the various planets and spacecraft have all been rotated into the same plane.

The observations discussed here were obtained on both Voyager spacecraft using identical plasma wave receivers which measure the electric field component of plasma and radio waves in the frequency range 10 Hz to 56 kHz by the use of 16-channel spectrum analysers and wideband receivers with passbands of 40 Hz to 12 kHz (ref. 1).

Observations

During the period from 30 August 1983 to the present, a very weak signal has been apparent in the 3.11-kHz channel of the Voyager 1 plasma wave instrument. The 3.11-kHz data for a 6-month interval are plotted as a function of time in Fig. 2. Each point in Fig. 2 represents a 51.2-min average of the electric field spectral density in the 3.11-kHz channel. It is apparent that after about day 242 (30 August) 1983 a smoothly varying signal was continuously present. Amplitude variations within the 51.2-min averaging intervals were small compared with the long-term variations seen in Fig. 2. Beginning at about day 292, the variations appear to be quasi-periodic with a period of about 26 days. This is close to the sidereal period of the Sun, hence, we suspect the solar wind may be responsible, in part, for the amplitude variations. There also seems to be a gradual overall variation in the amplitude of the emission with a rise from the onset near day 242 to the peak near day 285 and a subsequent

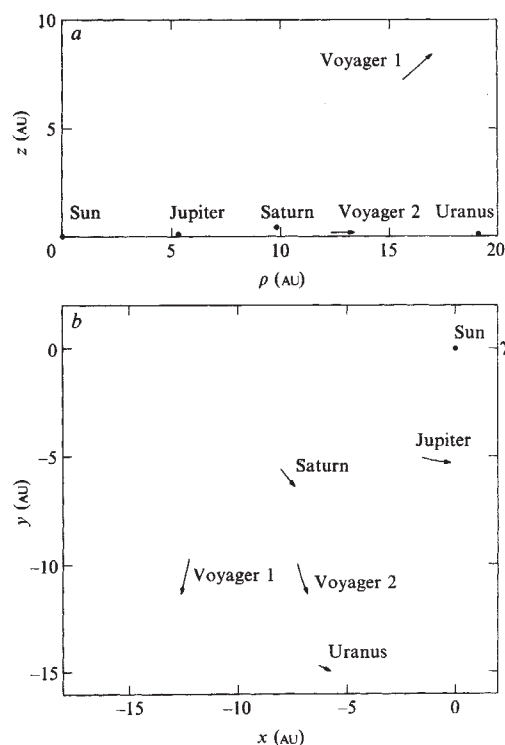


Fig. 1 The trajectories of Jupiter, Saturn, Uranus, and Voyagers 1 and 2 during the interval 30 August 1983 to 21 February 1984 when the Voyager 1 plasma wave receiver was detecting a new radio emission in the frequency range 2–3 kHz. Evidence from Voyager 2 indicates the radio emission may have been observable from that vantage point over a similar interval. *a*, The positions of the outer planets and the Voyager spacecraft rotated into a single meridional plane with the ρ -axis lying in the ecliptic plane. *b*, The projection of the trajectories of the same bodies in the ecliptic plane with the x -axis pointing in the direction of the first point in Aries, γ .

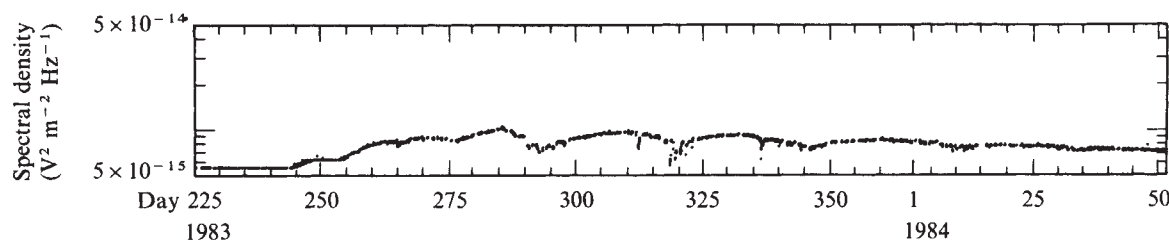


Fig. 2 Electric field spectral densities from the Voyager 1 3.11-kHz plasma wave receiver channel plotted as a function of time for slightly more than six months. Each point represents a 51.2-min average. The radio emission first appeared on day 243 1983.

decrease in amplitude towards the end of the plotted interval. During this entire time interval, no signal was detected in the 5.62-kHz channel. For a few one- or two-week periods a much weaker emission could be seen in the 1.78-kHz channel.

Several times during the interval plotted in Fig. 2, the wideband channel of the Voyager 1 plasma wave receiver was interrogated. For each wideband interrogation, four contiguous 48-s 'frames' of plasma wave data with high temporal and spectral resolution were obtained. In each wideband frame transmitted after day 242 1983, a weak band of emission near 3 kHz could be seen. Table 1 summarizes the wideband observations of the 3-kHz emission by both Voyager spacecraft. SCET is the spacecraft event time for each wideband observation and f_0 and $f_{\min}$ are the frequencies of the peak of the emission and the low-frequency cutoff, respectively. The electron plasma frequency as determined by the Plasma Science instrument (J. W. Belcher, personal communication) is labelled f_p .

An example of one of the wideband frames is shown in Fig. 3 in the format of a frequency-time spectrogram where the amplitudes of waves are plotted as a function of frequency and time. The most intense waves are black. In Fig. 3, the intense tones at frequencies below 1 kHz are interference from the operation of the onboard tape recorder used to record the wideband data. The very narrow tones at 2.4 and 4.8 kHz are the first and second harmonics of the spacecraft power supply.

The diffuse noise seen between 2 and 3 kHz is the emission which accounts for the signal received in the 3.11-kHz channel plotted in Fig. 2. This spectrogram is typical of the others received since day 242 1983, except for variations in the signal strength and minor variations in the lower frequency cutoff of the emission. The data from a 4-s interval in the spectrogram in Fig. 3 were averaged to form the spectrum shown in Fig. 4a which illustrates the detailed structure of the emission band. The power supply interference tone at 2.4 kHz, which rises out of the middle of the emission band, and a notch filter which serves to limit the power supply interference, cause some difficulty in the interpretation of the spectrum in Fig. 4. The gap

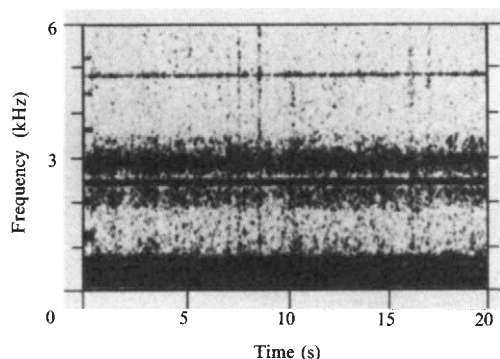


Fig. 3 Spectrogram showing the intensity of waves as a function of frequency and time for a 20-s interval. The diffuse emission seen between 2 and 3.5 kHz is the new radio emission reported here. The narrowband tones at 2.4 and 4.8 kHz as well as below about 800 Hz are spacecraft interference. Start time 2021: 52 SCET of day 307, 1983; $R = 17.9$ AU; celestial lat. = 25.5° .

in the spectrum at 2.4 kHz may be due to the notch filter, but we cannot rule out the possibility of a complex spectral structure which is inherent in the radio emission itself.

The Voyager 2 3.11-kHz channel sensitivity is not as great as that on Voyager 1 because of a partial failure in the Voyager 2 Flight Data System which occurred shortly after launch. Hence, at these low wave amplitudes we do not see any response in the Voyager 2 spectrum analyser channels. On the other hand, wideband frames obtained from Voyager 2 (which are unaffected by the Flight Data System failure) show a feature which is very similar to that shown in Fig. 3 (see Table 1). The earliest Voyager 2 wideband frame showing evidence of the emission was obtained on day 257 (14 September) 1983 when the spacecraft was 12.7 AU from the Sun. Samples showing evidence of the emission were also obtained on days 294, 307, and 339 of 1983 and day 3 of 1984. Because no wideband frames were obtained on Voyager 2 between day 228 (when the emission was not seen in the wideband data) and day 257, it is possible that the signal has been continuously present at Voyager 2 since about day 243 as was the case for Voyager 1. This close correspondence in onset times observed at two locations separated by nearly 10 AU suggests the two spacecraft are detecting a temporal change as opposed to a spatial variation.

Figure 4b shows a spectrum taken on day 307 by the Voyager 2 plasma wave instrument. The receiver threshold is not quite as low as in the case of Voyager 1, but the similarity between the emission just below 3 kHz with the feature shown in Fig. 4a can be readily seen. Note that the amplitude of the emission is nearly identical to the emission observed by Voyager 1.

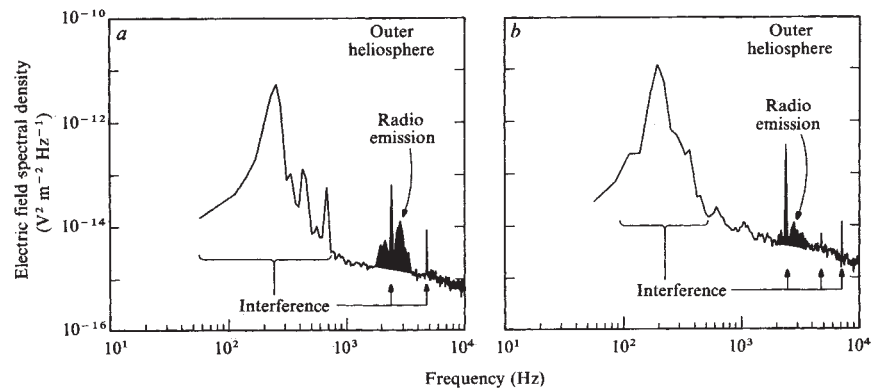
The Voyager 2 observations of the emission are important in our identification of the emission as a radio signal because the plasma instrument² on Voyager 2 can provide a measure of the local electron density, hence, plasma frequency. (The Voyager 1 plasma instrument has not functioned since shortly after the Saturn encounter.) If the emission is a freely propagating radio emission, then it must lie above both the electron plasma frequency and gyrofrequency. For four of the five days when wideband data are available, a determination of the density from the plasma instrument yielded values ranging from 0.02 to 0.04 cm^{-3} as shown in Table 1 (J. W. Belcher, personal communication). The respective plasma frequencies range from

Table 1 Wideband waveform observations of 3-kHz radio emissions

Year	Day	SCET	f_0 (kHz)	$f_{\min}$ (kHz)	f_p (kHz)
Voyager 1					
83	283	1407	2.8	1.9	NA
83	307	2022	2.9	1.8	NA
83	339	1720	3.0	1.9	NA
84	3	1619	3.0	2.1	NA
Voyager 2					
83	257	1940	3.0	2.7	1.4
83	284	1452	2.8	1.9	NA
83	307	1633	2.8	2.0	1.3
83	339	1747	3.0	2.6	1.3
84	3	1805	3.1	2.7	1.8

NA, not available.

Fig. 4 *a*, 4-s average spectrum taken from the spectrogram shown in Fig. 3 at 2021: 56 SCET. The shaded portion of the spectrum corresponds to the new emission. Note that the interpretation of the spectrum is complicated by the existence of a notch filter and power supply tone centred at 2.4 kHz. *b*, A 4-s average spectrum of the 3-kHz radio emission observed by Voyager 2 at a time (1632: 05 SCET) very close to when the emission shown in *a* was observed by Voyager 1. Note the similar amplitude of the features near 3 kHz in *a* and *b*.



1.3 kHz to ~ 1.8 kHz because $f_p(\text{Hz}) = 8,980\sqrt{n_e}$, where n_e is the electron density in cm^{-3} . For days 257, 307, and 339, the values for f_p are obtained from the 3-min determination of n_e made closest to the wideband frames and have errors of $\sim 10\%$. No reliable measurement is available for day 284 and the value for day 3 of 1984 is a daily average used in the absence of detailed analyses which are not yet completed. The low-frequency cutoff of the emission ranged from ~ 2 to nearly 3 kHz over the five intervals of wideband data. In each case, the emission exhibited a low-frequency cutoff which was well above the local electron plasma frequency. As the emission lies above f_p and the electron gyrofrequency in the solar wind is much less than f_p , the emission meets a necessary condition for being a freely propagating radio wave. It is easy to speculate that the periodic dips in the emission amplitude seen in the Voyager 1 data plotted in Fig. 2 are due to the passage of relatively high-density structures in the solar wind associated with high-speed streams which shield the spacecraft from the source of emission.

With the available data, it is not possible to confirm that the emission is a freely propagating radio wave; however, we can eliminate local plasma phenomena with a high level of confidence. There are two types of local (non-propagating) phenomena which have been observed at, or above, the local plasma frequency in the solar wind. The most familiar waves in this frequency regime are electron plasma oscillations or Langmuir waves. The observed waves are almost certainly not plasma oscillations as the bandwidth of the observed emission is much larger than that for plasma oscillations, particularly at low amplitudes and plasma oscillations are characterized by a very sporadic temporal character. The observed emission shows only very smooth variations in amplitude with time.

The second type of local phenomena which might be considered are quasi-thermal electrostatic plasma waves³. Hoang *et al.*³ discuss the spectral form of the thermal plasma emission in detail for both long and short (compared with the Debye length) antennas. For the typical solar wind observed in this epoch by the two Voyagers, the density is 0.05 cm^{-3} , hence, even with a temperature as low as 1 eV, the Debye length is > 30 m, compared with an effective antenna length on the two Voyagers of 7 m. Therefore, the short antenna results of Hoang *et al.* are relevant to this discussion. The quasi-thermal plasma noise detected by a short antenna has a power law spectrum with an index of -1.5 and shows no peak and no cutoff at f_p . The spectra shown in Figs 3 and 4 exhibit both a low-frequency cutoff and at least one peak. These waves, therefore, are clearly not the thermal electrostatic waves discussed by Hoang *et al.*³.

Table 1 also shows that the frequency of the peak in the emission is nearly the same at both spacecraft despite the relatively large variations in local plasma conditions which would be expected between two points separated by ~ 10 AU. The nearly constant frequency would not be expected if the emission were a local effect. Finally, the fact that Voyagers 1 and 2 both began detecting the 3-kHz emission at about the same time despite being nearly 10 AU apart argues strongly for a freely

propagating radio emission and against a local plasma phenomenon.

Possible sources

Several possible sources exist in, and beyond, the outer heliosphere for the radio emission reported here. The emission is most reminiscent of continuum radiation trapped in Jupiter's distant magnetotail because it lies in a very similar frequency range. Figure 5 shows an example of the continuum radiation observed $2,019 R_J$ downstream from Jupiter⁴, where R_J is the radius of Jupiter. Given that the continuum radiation is confined to the tail cavity by the surrounding high-density solar wind close to Jupiter and also given the geometry of the observations presented in Fig. 1, it is clear that the emissions would have to propagate down the jovian tail to a point where the waves are at a frequency greater than the surrounding solar wind plasma frequency. Having reached this 'window' they could subsequently propagate freely and possibly escape in the direction of the Voyager spacecraft.

However, a comparison of the two spectra from the outer heliosphere in Fig. 4 with that in Fig. 5 taken in Jupiter's tail suggests that the two emissions have different sources. The spectrum of trapped continuum radiation varies as $f^{-\alpha}$ where α ranges between 3 and 6 (ref. 5). The radiation in Fig. 4 clearly does not show that type of frequency dependence. Moreover, the other waveform samples from both Voyagers of the new radio emission do not deviate greatly from the form shown in Fig. 4. Hence, to identify the spectra in Fig. 4 with jovian continuum radiation, one would have to rely on the solar wind medium to reshape the spectrum. The general consistency of

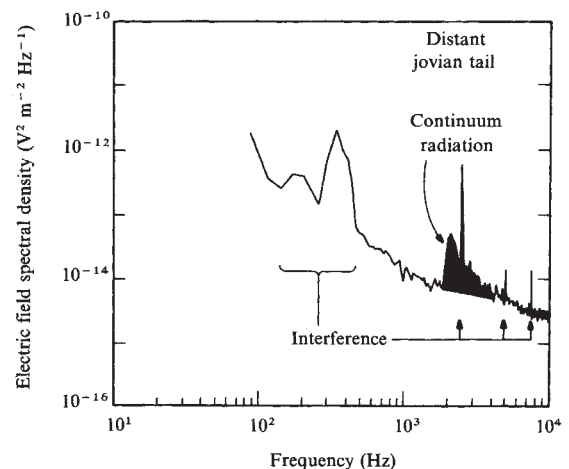


Fig. 5 A 4-s average spectrum of continuum radiation trapped in the jovian tail $2,019 R_J$ downstream from Jupiter (shaded portion) obtained by Voyager 2 on 1 February 1980 at 0016: 26 SCET. Notice that this emission is sharply peaked at lower frequencies in contrast to the emission shown in Fig. 4.

the recently obtained spectra is inconsistent with such a shaping process.

An analysis of the amplitude of the new emission also argues against a jovian source. If the continuum radiation escaping from the tail of Jupiter were to be emitted from some point along the tail (assumed to extend in the anti-solar direction from Jupiter), one would expect to see an approximate R^{-2} dependence in the intensity of the emission from the exit 'window'. From Fig. 1, we see that Voyager 1 and 2 are about 14.8 and 6.4 AU from the jovian tail axis, respectively; hence, one would expect a signal more than a factor of 5 times larger at Voyager 2 than at Voyager 1, but waveform samples taken on both spacecraft on 3 November (shown in Fig. 4) indicate that the spectral density at the peak of the emission band is nearly identical at both spacecraft. (We have used a calibration technique similar to that used by W.S.K. *et al.*⁵ which avoids the Voyager 2 calibration uncertainty.) Also, the intensity of the Voyager 1 spectrum obtained at 18 AU is only a factor of about 3 less than that obtained only 2,019 R_J (~ 1 AU) downstream from Jupiter. Hence, it seems difficult to understand the emission as jovian in origin on the basis of intensity.

We have also performed a power spectrum analysis on the data presented in Fig. 2 to look for variations near the rotation period of Jupiter since some of the radiation detected in the distant jovian tail shows amplitude fluctuations with a period of about 10 h (ref. 6). No periodicity near 10 h was detected.

Thus, while Jupiter initially seems a likely source for the emission on the basis of the frequency of the emission, the spectral shape, the lack of spatial variation in amplitude, and the lack of 10-h periodicities cast serious doubts on this interpretation. Below, we examine other possibilities for the source of the emission.

Continuum radiation trapped within Saturn's magnetosphere has also been detected⁵. Although the position of Voyager 1 might seem to be ideal to detect such saturnian radiation when one examines only the ecliptic plane projection in Fig. 1b, Fig. 1a shows that the spacecraft was actually 8 AU north of the ecliptic; hence, it is difficult to understand how any radiation from Saturn could propagate to either of the spacecraft. Moreover, as with Jupiter, the saturnian continuum radiation spectrum does not match the spectra in Fig. 4. In addition, the saturnian radiation is even weaker than Jupiter's, thus, we conclude that Saturn is not a likely source for the newly detected radio emission.

Uranus is another possible source for the radiation as we do not know what radio emissions from Uranus might look like. Using observations of the Earth, Jupiter, and Saturn as guides, the upstream planetary radio emission which most closely resembles the new emission is escaping continuum radiation from the Earth^{7,8} and Jupiter⁹. The escaping continuum radiation from those planets is extremely weak, however, and even if Uranus had similar emissions, they could not be detected by either Voyager spacecraft at their present distances unless the source was extremely bright. Other types of planetary radio emissions are intrinsically more intense than the continuum radiation; their dynamic spectra, however, are more highly structured than the emissions reported here. Nevertheless, we cannot rule out Uranus as a source of the radio noise reported above until the Voyager 2 encounter with that planet in 1986, although, we now feel that Uranus is not a likely source.

It is possible that the new radio emission detected by Voyagers 1 and 2 is generated outside the heliosphere. With our lack of information about the region, particularly at such low frequencies, any number of sources might be contemplated. Once such source is a fast pulsar which might emit enough energy at low harmonics of its pulsing rate to be detected in the frequency range of a few kHz. However, we suggest that the broad bandwidth of the newly detected signal is sufficient to rule out pulsars as the source. In addition, the power flux of the band shown in Fig. 4a integrated over the 1-kHz bandwidth is of the order of 10^{-14} W m⁻². The radio luminosity of the Crab Nebula integrated over a bandwidth of 10^8 – 10^9 Hz is about 10^{31} erg s⁻¹

(ref. 10), hence, at a distance of ~ 1 kpc this corresponds to 10^{-16} W m⁻², much less than the value observed by the Voyager spacecraft. (An equivalent, but alternate comparison, is to calculate the power spectral density from the Crab assuming the 10^{31} erg s⁻¹ is emitted in a 1-kHz bandwidth. This would be about 10^{-19} W m⁻² Hz⁻¹ compared with a value of 10^{-17} W m⁻² Hz⁻¹ observed by Voyager.) It is interesting to think of what other sources of radio noise lie beyond the Solar System, but perhaps fruitless to speculate on them at this point.

We have considered one other possibility for the source of the radio emission: the heliopause, itself. It is reasonable to assume that there is a supersonic shock marking the heliopause (sometimes referred to as the terminal shock) and shocks are known to be the source of weak, relatively narrowbanded radio noise. Specifically, Dunckel¹¹, Gurnett¹² and Hoang *et al.*¹³ have all reported radio emissions at $2f_p$ emanating from the vicinity of the Earth's bow shock and a similar phenomenon has been reported near an interplanetary shock⁴. Theories for the emission at $2f_p$ have been offered by Fung *et al.*¹⁴ and Cairns and Melrose¹⁵. The fact that this type of emission is quite weak is countered by the large source size of the heliopause and resulting large solid angle as seen by an observer in the outer heliosphere. The smoothly varying temporal character and bandwidth of the new emission are consistent with the $2f_p$ model.

If, indeed, the heliopause is generating the radio emission at $2f_p$, some interesting statements can be made concerning the size of the heliosphere. This scenario implies that the plasma frequency at the heliopause is half of 3 kHz, or 1.5 kHz; n_e at the heliopause is then ~ 0.03 cm⁻³. One can expect a maximum jump in the density at the shock of a factor of 4 (and we also assume the interstellar density is higher than just inside the heliopause), meaning the density just inside the heliopause could be as low as 0.008 cm⁻³. This corresponds to a plasma frequency of about 780 Hz. Because the density in the solar wind falls as $1/R^2$ (assuming constant solar wind speed) and the plasma frequency is proportional to $\sqrt{n_e}$, the plasma frequency varies as $1/R$. Given that f_p is ~ 2 kHz at 18 AU (using Fig. 3) and that it drops to 780 Hz at the heliopause, it is easy to solve for the radial distance of the heliopause, that is ~ 46 AU.

Obviously, the reasoning above assumes no variations in the solar wind density as a function of time and is subject to several simplifying assumptions. Also, if one assumes the density in the interstellar medium is less than just inside the heliopause, the minimum f_p in the heliosphere is 1.5 kHz and Voyager 1 is very close to that point. Taking the value of 46 AU (with liberal error bars), however, we can compare it with other predictions of the distance to the heliopause. One popular method is to interpret the cosmic ray gradient to arrive at a distance to the boundary. One such prediction of >65 AU by Webber and Lockwood¹⁶ was based primarily on Pioneer 10 observations. As Voyager 1 is travelling more or less antiparallel to the interstellar wind flow and Pioneer 10 is headed in nearly the opposite direction, it is commonly assumed that the heliopause will be encountered at smaller heliocentric radial distances by Voyager 1, hence, 46 AU does not seem to be unreasonable. Voyager 1 will be at a distance of about 46 AU in 1991 and perhaps only then will we be certain of the source of this new radio emission.

Conclusion

Voyager 1 and 2 observations of a very weak signal in the range 2–3 kHz in the outer heliosphere at distances of 13–19 AU appear to be evidence of a newly discovered radio emission. While the frequency of the radio emission is similar to continuum radiation in the distant jovian tail, other observations are inconsistent with Jupiter as a source. One other possible candidate for the emission source is $2f_p$ radiation from the heliopause. This is a most exciting possibility, because if this is the case, these observations are our first glimpse of the edge of the heliosphere.

We thank J. W. Belcher for information on the electron density observed by the Voyager 2 plasma instrument during some of the events used in this study. We also acknowledge the program-

ming support of L. J. Granroth. Research at The University of Iowa was supported by NASA through contract 954013 with the Jet Propulsion Laboratory and grant NAGW-386 with NASA Headquarters. Research at TRW was supported by NASA

through contract 954012 with the Jet Propulsion Laboratory. The research described above was performed in part at the Jet Propulsion Laboratory, California Institute of Technology and was jointly sponsored by The University of Iowa and NASA.

Received 14 May; accepted 22 August 1984.

1. Scarf, F. L. & Gurnett, D. A. *Space Sci. Rev.* **21**, 289-308 (1977).
2. Bridge, H. S. *et al. Space Sci. Rev.* **21**, 259-287 (1977).
3. Hoang, S. *et al. J. geophys. Res.* **85**, 3419-3430 (1980).
4. Kurth, W. S., Gurnett, D. A., Scarf, F. L., Poynter, R. L. & Sullivan, J. D. *J. geophys. Res.* **86**, 8402-8412 (1981).
5. Kurth, W. S., Scarf, F. L., Sullivan, J. D. & Gurnett, D. A. *Geophys. Res. Lett.* **9**, 889-892 (1982).
6. Kurth, W. S. *et al. J. geophys. Res.* **87**, 10,373-10,383 (1982).
7. Gurnett, D. A. & Shaw, R. R. *J. geophys. Res.* **78**, 8136-8149 (1973).

8. Kurth, W. S., Gurnett, D. A. & Anderson, R. R. *J. geophys. Res.* **86**, 5519-5531 (1981).
9. Gurnett, D. A., Kurth, W. S. & Scarf, F. L. *Nature* **302**, 385-388 (1983).
10. Pacini, F. in *The Crab Nebula* (eds Davies, R. D. & Smith, F. G.) 394-406 (Reidel, Dordrecht, 1971).
11. Duncel, N. thesis, Stanford Univ., California (1974).
12. Gurnett, D. A. *J. geophys. Res.* **80**, 2751-2763 (1975).
13. Hoang, S., Fainberg, J., Steinberg, J. L., Stone, R. G. & Zwickl, R. H. *J. geophys. Res.* **86**, 4531-4536 (1981).
14. Fung, S. F., Papadopoulos, K. & Wu, C. S. *J. geophys. Res.* **87**, 8077-8080 (1982).
15. Cairns, I. H. & Melrose, D. B. *J. geophys. Res.* (submitted).
16. Webber, W. R. & Lockwood, J. A. *J. geophys. Res.* **86**, 11,458-11,462 (1981).

A third type of murine T-cell receptor gene

Yueh-hsiu Chien, Daniel M. Becker, Tullia Lindsten, Maki Okamura, David I. Cohen* & Mark M. Davis

Department of Medical Microbiology, Stanford University School of Medicine, Stanford, California 94305, USA

By subtractive cDNA hybridizations, we have isolated a new species of T-cell receptor cDNA clone whose predicted amino acid sequence has homology to variable, constant, joining and diversity segments of immunoglobulins and T-cell receptors. The corresponding genomic sequence is also rearranged in several T-cell DNAs. The four potential N-linked glycosylation sites, frequency of expression and predicted molecular weight (27,800) of this molecule make it a likely candidate for the α -chain of the T-cell receptor. Expression data also indicate that this gene may be activated at a later stage of T-cell differentiation than the β -chain.

RECOGNITION of foreign entities by T cells is strikingly similar to that of B cells in the large number of specificities that are engendered and in the fixation in particular cell lineages of these specificities (reviewed in ref. 1). Unlike the mediators of B-cell recognition (immunoglobulins), however, T-cell receptors seem able to 'recognize' antigen only in conjunction with self-major histocompatibility complex (MHC) determinants. Cytotoxic T cells (T_c) typically require class I antigenic determinants ($H-2$)²⁻⁴ and helper T cells (T_H) require class II (Ia) determinants⁵. The affinity for self-MHC seems to be 'learned' in some way in the thymus^{6,7} before exposure to antigen and apparently allows T cells to focus on abnormalities of the organism's own cells. In addition, helper T cells up-regulate and suppressor T cells down-regulate B cells and other immune system-related responses.

Uncertainties concerning the biochemical and genetic nature of the T-cell receptor molecule(s) has been a major stumbling block in understanding the phenomena described above. Continual progress, especially in the past few years, has begun to provide some answers. In particular, T-cell cloning methodologies using either long-term lines or hybridomas¹ have allowed a better definition of the antigen-MHC reaction phenomena and have also provided large quantities of homogeneous cell populations of a single, defined specificity. This has also led to the development of clone-specific (anti-idiotypic) antisera⁸ and monoclonal antibodies⁹⁻¹³ which can either block or stimulate the response of the target cells to antigen¹⁰⁻¹³. These 'clonotypic' monoclonal antibodies were used successfully to precipitate the putative T-cell receptor molecule itself, an 80,000-90,000 dalton molecular weight (80-90K) sulphhydryl-linked heterodimer⁹⁻¹³ which on reduction yields α - and β -subunits of 40-50K each. These subunits seem to have both variable and constant fragments in tryptic peptide maps^{14,15}. Recently, cDNA clones encoding one chain of the T-cell receptor were isolated in mouse^{16,17} and man¹⁸. N-terminal

amino acid sequencing established that these clones encode the β -chain¹⁹ of the T-cell receptor heterodimer. The β -chain cDNA clones encode a 32K protein having variable, constant and joining (J) region elements remarkably similar to those of immunoglobulins^{17,18}. More recently, D (diversity) elements have also been shown to play a part²⁰⁻²³ in the formation of the

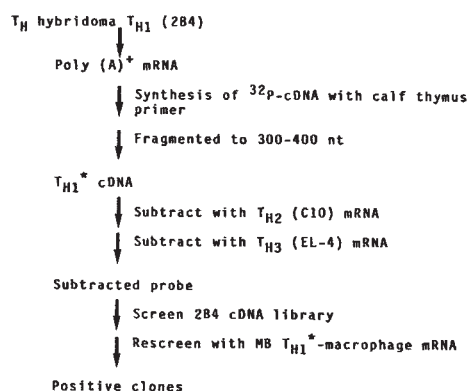


Fig. 1 cDNA subtractions and screening strategy. Calf thymus DNA was used to synthesize randomly primed ³²P-labelled cDNA by standard procedures⁴⁷ from poly(A)⁺ cytoplasmic 2B4 mRNA. The cDNA was initially 700 nucleotides (nt) average length and was allowed to fragment by autoradiolysis to ~300 nt long over a 2-week period. The subtractive hybridization was carried out by procedures described previously⁴². The ³²P-cDNA was first hybridized and hydroxyapatite-selected with the T_H hybridoma C10 mRNA, then with the T_H -like lymphoma line, EL-4. The twice-subtracted probe was then hybridized to a filter of 2B4 cDNA library in the vector λ gt10 as described elsewhere²⁰. Approximately 20,000 plaques were screened. Seven positive plaques were picked and rescreened with a probe of oligo(dT)-primed cDNA made of membrane-bound polysomal mRNA from 2B4, subtracted with mRNA from the P388D₁ macrophage line (MBT_H-Mac).

* Present address: Laboratory of Immunology, NIAID, National Institutes of Health, Bethesda, Maryland 20205, USA.

variable-region exon of this gene. The mechanism of rearrangement seems very similar to that of immunoglobulins²⁰⁻²², the only major difference being the extreme diversity and apparently accelerated rate of divergence of the different variable regions, apparently due to the presence of additional hypervariable regions in the molecule²⁴. These data have helped to elucidate the problem of diversity and clonal fixation of specificities by direct analogy with immunoglobulin.

The other questions specific to T-cell recognition are more difficult and require knowledge (and possession) of the gene encoding the second chain of the heterodimer (α). An apparent

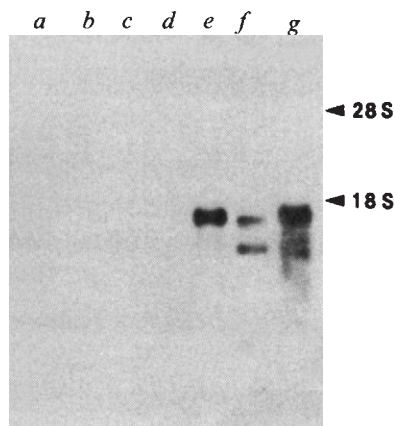


Fig. 2 T cell-specific expression of TT11. The TT11 cDNA clone was labelled by nick-translation⁴⁵ and hybridized to a Northern blot⁴⁵ containing a panel of mRNAs as follows: *a*, Bal 17, B-cell lymphoma; *b*, M104E, plasmacytoma; *c*, 3T3, fibroblast line; *d*, P388D₁, macrophage line; *e*, 2B4; *f*, EL-4; *g*, BW5147. All are poly(A)⁺ cytoplasmic RNAs prepared by standard procedures⁴⁵.

answer in this regard was obtained by Saito *et al.*²⁵, who used subtractive cDNA selection methodologies¹⁶ to obtain a second immunoglobulin-like, T cell-specific cDNA clone (HDS4 and related species) which is rearranged in T_c DNAs and has elements homologous to variable, constant and *J* regions. Although this gene may encode one of the chains of the T-cell receptor for murine cytotoxic T cells, it seems unlikely that it also encodes the α -chain for the murine helper T cells for the following reasons: (1) there are no N-linked glycosylation sites (N-X-S_T) in the amino acid sequences predicted for the HDS4 series

clones, whereas three N-linked additions have been found in α -chains from three different murine helper T-cell lines²⁶⁻²⁸, including 2B4 (ref. 26), the subject of this report. (2) The HDS4 gene is not expressed in many T_H-cell lines (our results and S. Tonegawa, personal communication). That T_H and T_c cells use the same β -chain locus has been established^{25,29-31} by molecular genetic analysis. (3) A consistently low level of expression has been seen for this gene with respect to β -chains, that is, 1% of the level of β -chain expression in thymus and 2% of the level in spleen cDNA libraries (this paper), and $\leq 10\%$ of the β -chain level in some T_c-cell libraries (Z. Eshhar, personal communication). Although differences in translational efficiency or stability are possible, all known multi-chain molecules have mRNA species present in approximately equal amounts (see below).

In view of these questions, we generated a new cDNA probe designed to detect variable regions which might differ between various T-cell lines, and used this to screen a T_H cDNA library from the cytochrome *c*-specific hybridoma, 2B4 (ref. 12). One of the cDNA clones obtained in the screen, TT11, appears to encode a protein with many of the expected features of an α -chain and is expressed in both T_H and T_c cells. If this is confirmed to be the case, its characterization will identify the missing portion of the T-cell receptor molecule. However, this result also raises the interesting question of why there are three (or more) immunoglobulin-like chains in at least some T cells.

Isolation of cDNA clones

As peptide mapping has shown that the α -chains of the T-cell receptor molecules isolated from T_H cells are as diverse as the β -chains¹⁵, and given the distinct immunoglobulin-like structure of the β -chain of the T-cell receptor^{17,18} and the likelihood that the α -chain is encoded by a similar molecule, it seemed feasible to prepare a variable (*V*)-region-specific subtracted cDNA probe between T cells of differing specificities. This was achieved as shown in Fig. 1, where randomly primed, labelled cDNA⁴⁷ from the mRNA of the helper hybridoma, 2B4 (ref. 12), was synthesized; this probe was allowed to fragment by autoradiolysis to an average size of 300–400 nucleotides and then subtracted, first with C10 mRNA, then with EL-4 mRNA. The resultant probe should be enriched in 2B4 *V*-region-specific sequences, and was then used to screen a 2B4 cDNA library made in the vector λ gt10²⁰. The use of calf thymus DNA as a random primer gives fairly equal reverse transcription over the entire length of each mRNA species, thus eliminating the 3'-end bias of oligo(dT) priming. The size of 300–400 nucleotides is optimal in that it is large enough to give good reactivity in hybridizations,

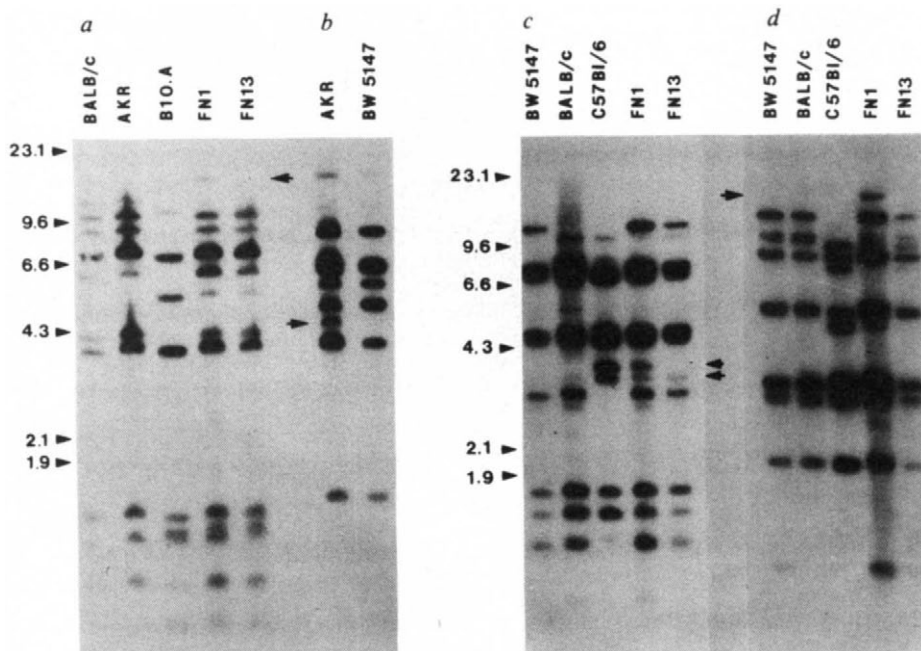


Fig. 3 Rearrangement of the TT11 gene. Genomic DNAs from the livers of different mouse strains, various T-cell lines and hybrids of the B-cell lymphoma L10A, were digested with: *a*, HindIII; *b*, EcoRV; *c*, XbaI; *d*, BglII and electrophoresed through 0.7% agarose, blotted onto nitrocellulose and hybridized by standard methods⁴⁵ to a probe from the 5' half of TT11 (RI-RV; see Fig. 4). Arrows indicate altered bands, arrowheads indicate molecular weight markers. In the HindIII digest (*a*) the B10.A pattern is identical to that of the true parental, C57Bl/6.

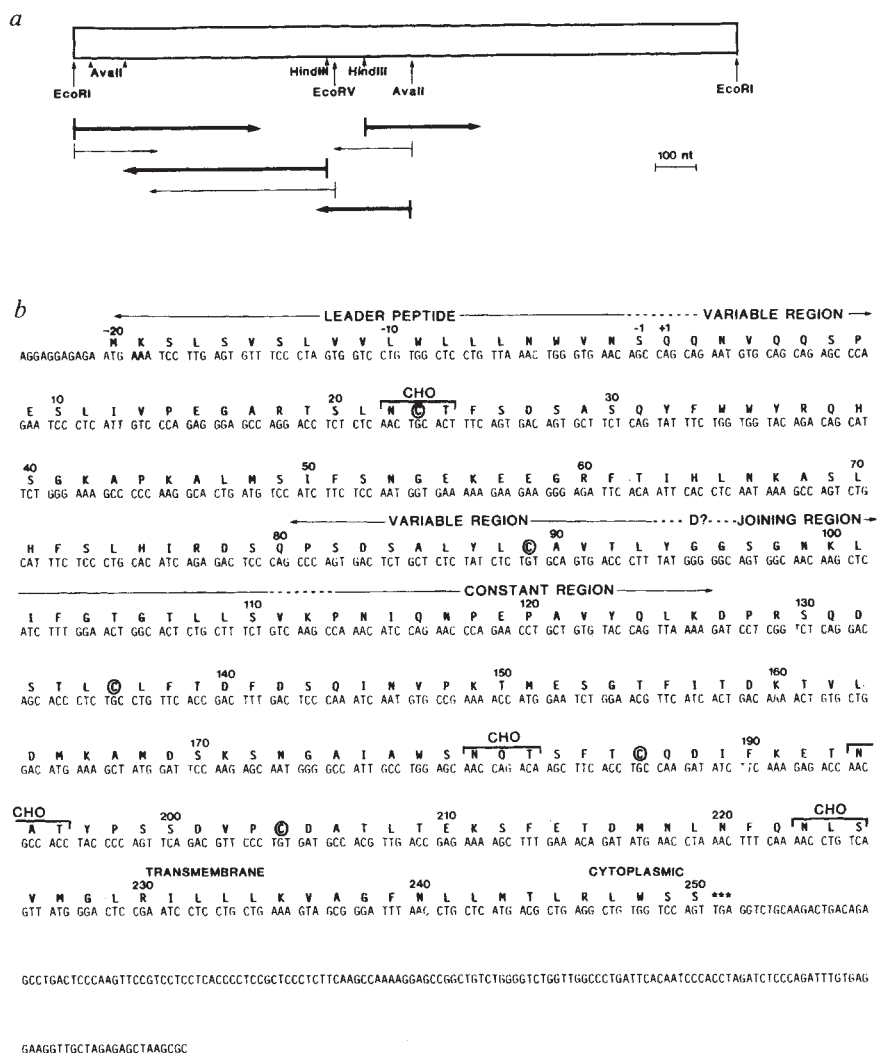


Fig. 4 *a*, Sequencing strategy for the coding portion of TT11 cDNA clones. The *EcoRI* sites are generated by the RI linkers. Thin lines indicate 5'-end labelling with polynucleotide kinase and thick lines indicate 3'-end labelling with the large (Klenow) fragment of DNA polymerase I. *b*, The determined nucleotide sequence and the predicted amino acid sequence. The numbers above the amino acid sequence designate the residue positions. The leader, variable, possible D, joining, constant, transmembrane and cytoplasmic regions are indicated although the exact boundaries are somewhat ambiguous. Cysteines are circled. Potential signal sequences for N-linked glycosylation are indicated by CHO.

yet, as it is approximately the same size as immunoglobulin and T-cell receptor β -chain V regions, at least some fractions of the species will not have C region-hybridizable sequences. One complication in such a screening protocol is the fact that even phenotypically identical cell lines and hybridomas appear to differ by a small but significant fraction of their gene expression (0.2–0.4%; M.M.D., S. M. Hedrick and W. E. Paul, unpublished observation), perhaps through random gene loss or activation. As β -chain mRNA expression is generally much lower than this in hybridomas (0.02%)²⁰, it was necessary to subtract the probe with two separate mRNAs (as shown in Fig. 1) from T cells which do not share the same specificity. C10 and EL-4 have distinct T_H phenotypes similar to 2B4 but are not cytochrome *c*-reactive. There is, of course, some uncertainty in constructing such strategies (see Fig. 1) because even with cells of differing specificity, the same or cross-reactive V regions may be expressed. Seven positive plaques obtained in this primary screen were rescreened with another cDNA probe made from the membrane-bound polysomal RNA of 2B4 subtracted with RNA from the macrophage line, P388D₁ (ref. 32). This provides an additional selection for T-cell receptor molecules^{16,25}. Three of the seven plaques were positive with the MBT_H-Mac probe (using the convention of ref. 16) and two of these three cross-hybridized with each other. The one with the largest insert, TT11, was analysed further.

Pattern of expression and rearrangement

Figure 2 shows that, on Northern blot analysis, TT11 is T cell-specific, as it is not expressed in the B-cell lymphoma (Bal 17), a plasma cell line (M104e), a macrophage line

(P388D₁) or a fibroblast line (3T3). It is expressed in all T cells currently surveyed, including the two, used in the T*–T subtractions: EL-4 (Fig. 2) and C10 (data not shown), as well as in three T_c-cell lines (T.L. *et al.*, work in progress). It is interesting to note the two distinct bands at 1.8 and 1.3 kilobases (kb) in the EL-4 lane; the possible significance of this with respect to D-like regions in this molecule will be discussed below.

The gene(s) encoding the TT11 cDNA clone show distinct signs of rearrangement (Fig. 3), characteristic of all the immunoglobulin-like B-cell-(IgH, Ig κ and Ig λ)³³ and T-cell-(β , HDS4)^{16,25} derived genes. Arrows in Fig. 3 indicate bands which have changed position with respect to the parental genomes in the T-cell DNAs. FN1 and FN13 are keyhole-limpet haemocyanin (KLH)-reactive T_H hybridomas (referred to in ref. 34 as FN1–18 and FN13–21) derived, respectively, from BALB/c $\times$ C57B/6 strain mice and the AKR strain thymoma line, BW5147. In Fig. 3a, a new band appears in a *HindIII* digest of FN1 (arrow) with respect to parental DNAs. In Fig. 3b, one band in an *EcoRV* digest of AKR liver DNA disappears in BW5147, and an *EcoRI* digest shows a new band appearing in BW compared with the parental liver DNA (data not shown). In Fig. 3c, the arrows indicate two bands which are polymorphic for an *XbaI* digest of C57B/6 DNA; both of these bands are present in the FN1 hybrid but only one occurs in FN13, which could only be due to a rearrangement or partial deletion of the chromosome. Figure 3d shows a new band in FN1 versus the parental DNA samples in a *BglII* digest. Although no digest shows evidence of rearrangement in all T-cell DNAs, these results provide sufficient indication of such events to suggest that TT11 is another T-cell receptor-like gene.

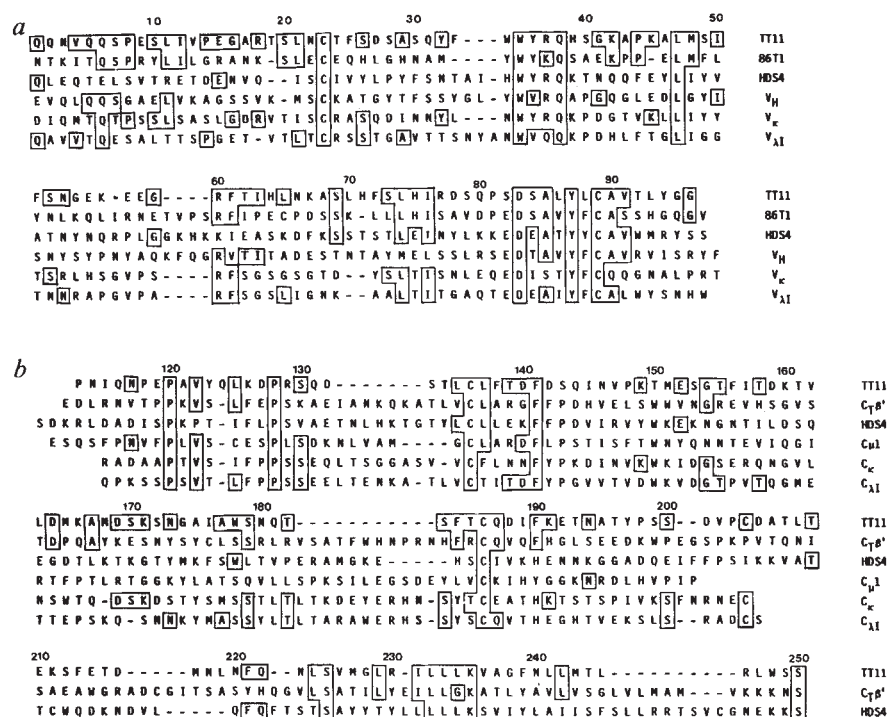


Fig. 5 Similarities of the TT11 amino acid sequence to variable (a) and constant (b) regions of a murine T-cell receptor β -chain 86-T1, HDS4 and immunoglobulin. Similarities are boxed, with gaps represented by dashes (-).

TT11 homology with immunoglobulins

The TT11 cDNA clone was partially sequenced by the procedure of Maxam and Gilbert³⁵ using the strategy shown in Fig. 4a; the sequence is shown in Fig. 4b. The clone was oriented by a poly(A) stretch (~150 nucleotides) at the 3' end and sequencing of the 5' half revealed a long open reading frame of 810 nucleotides, with an initiation codon (ATG) within the first 12 nucleotides (Fig. 4b). The sequence has regions similar to the immunoglobulin leader, variable regions, J region and constant region, as is the case also for the T-cell receptor β -chain^{17,18} and the HDS4 gene clone of Saito *et al.* Also, as in the latter two T cell-specific genes, TT11 has an extra cysteine outside the putative intra-domain cysteines in the constant region (which probably bonds to another chain), followed by a transmembrane portion and a potential cytoplasmic region, all of which are encoded as separate exons in the β -chain genes^{36,37}. There are four potential N-linked glycosylation sites, as indicated in Fig. 4b, similar to the four or five found in different β -chain sequences^{17,20}. Probably only three of the four potential sites are available for glycosylation as the most carboxy-terminal one is embedded in the putative transmembrane region (Fig. 4b).

The exact position at which the murine T-cell receptor protein is processed is unknown, but by analogy with all of the immunoglobulins (H, λ , κ), it should occur just before the glutamine at position +1 (see Fig. 4b). This is consistent with the finding that murine α -chains are blocked at their N-terminus (J. Allison, personal communication, and L. Samelson, personal communication). Alternatively, based on the N-terminal amino acid sequence of the human β -chain from the REX T-cell line¹⁹, the processing point would be just before the asparagine at position +3. In the first case, the molecular weight would be 27.8K, in the second slightly less; this agrees well with the data of Allison and colleagues²⁷ who obtained a molecular weight of 27K from a murine α -chain stripped of N-linked sugars with endoglycosidase F.

Comparisons with other putative T-cell receptor molecules and immunoglobulin V, C and J regions are shown in Figs 5 and 6, respectively. While the overall homology of the putative V and C regions with their immunoglobulin counterparts (10–26%) is relatively low, many of the conserved residues found

in all of five known immunoglobulin-like genes are present, particularly in the V region (Fig. 5) and J-region elements (Fig. 6). The cysteines in the TT11 variable region are 65 amino acids apart, which is identical to the spacing in the light chains, and the sequence WYRQ starting at residue 35 and DSA-Y-CAV at residues 83–91 are also highly conserved in most immunoglobulin and T-cell receptor V regions. As with the β -chain and HDS4, the putative J region is the most highly conserved portion, with 7 out of 16 residues homologous with the β -chain consensus sequence³⁶ and 9 out of 16 the same as J_T3 (ref. 36). Whereas the constant region has the least homology of any of these (10–15%), its overall structure is very similar to that of β -chain and HDS4. Somewhat more homology (up to 259) may be obtained in the constant region by using the cysteine at position 204 as the intradomain linkage, but this complicates matters with respect to the presumptive β -chain linkage.

Also characteristic of the T-cell receptor sequences, and shared with TT11, is the sequence ILLK in the putative transmembrane region. The conservation of a charged amino acid (lysine or K) in a transmembrane region is unusual and is not found in other members of the immunoglobulin superfamily; its presence in all three of the putative T-cell receptor molecules may indicate that some unusual multi-chain structure is formed with other molecules in the membrane.

Table 1 Frequencies of T-cell receptor expression

	C _T β	TT11	HDS4
Thymus	0.2%	0.01%	0.002%
Con A stimulated spleen	0.1%	0.03%	0.002%
2B4	0.02%	0.03%	<0.001%

Approximately 200,000 phage from each library were screened by standard procedures in duplicate. The C_T β probe was 86T5 (ref. 17), a DJC²²-containing cDNA clone; the HDS4 probe was a 150-nucleotide synthetic polymer corresponding to the constant region of HDS4 from a.a. 120→a.a. 180. The former was labelled by nick-translation and the latter by a fill-in reaction with Klenow polymerase. The TT11 probe was labelled by nick-translation.



Fig. 6 Amino acid homologies among the putative J region of TT11 and the J-region sequences of a murine T-cell receptor β -chain, J_T3 , HDS4 and immunoglobulins. Similarities are boxed, with variable residues represented by dashes (-).

Possible D regions

Another interesting feature of the TT11 sequence analysis is the strong indication that D regions are present in this molecule just as in the β -chain of the T-cell receptor²⁰⁻²³ and immunoglobulin heavy-chain genes³³. In particular, just 5' to the SGN amino acid sequence of the putative J region (which marks the 5' border of J_T3 in the β -gene complex), there is the nucleotide sequence GGGGG. This sequence occurs at the point where a D-region element would begin in the β -chain and is characteristic of the β -gene D regions (7 out of 14 contain runs of between 3 and 7 Gs on their 3' side)³⁸ as well as many immunoglobulin heavy-chain D regions³³. Although postulated to be the result of enzymatic addition in immunoglobulins⁴⁸, at least in some β -chain and IgH D regions, these poly(G) sequences are present in germ-line sequences^{22,23}. None of the immunoglobulin light-chain V-region genes sequenced has such poly(G) sequences before the J element junction³⁹. The Northern blot data of Fig. 2, where two bands are clearly visible in EL-4 (and other T_H lines, data not shown), also support the existence of independently assorting D gene segments. This two-band pattern is characteristic of the DJC transcript of the β -chain which is 300 nucleotides shorter than the VDJC transcript^{22,23}. Taken together, these data suggest that the TT11 gene will have D segments, adding to the potential diversity.

Frequency of expression

One characteristic of mRNA species which contributes to a multimeric structure is that they appear to be present in equal or nearly equal percentages in the steady-state mRNA population in cases that are known. Immunoglobulin heavy and light chains⁴⁰, Ia α - and β -subunits⁴¹, H-2 and β_2 -microglobulin⁴², and α - and β -globins⁴³ all follow this rule to within a factor of ≈ 2 . Although differential translation efficiencies or turnover rates could conceivably produce large differences in subunit mRNA frequencies, they do not seem to do so in these cases. To compare the steady-state mRNA frequencies in a relatively quantitative manner, we surveyed thymocyte, concanavalin A (Con A)-stimulated spleen and 2B4 cDNA libraries with TT11, HDS4 and $C_T\beta$ probes. The results (Table 1) indicate that TT11 and $C_T\beta$ are present in fairly similar frequencies (1:1-1:3) in the 2B4 and Con A spleen libraries, whereas HDS4 is much rarer, being present in $\frac{1}{50}$ th the amount of β -chain in the Con A spleen library and absent from the 2B4 library. This equivalence in frequency, together with the other data presented here, makes TT11 a more attractive candidate for the α -chain of the receptor in most T cells. Also of interest from the data in Table 1 is the change in the ratio of TT11 to β -chain in immature versus mature T cells. Specifically, the TT11/ β -chain ratio changes from 1:20 in thymocytes to 1:3 in Con A-stimulated spleen cells, suggesting that TT11 gene expression (and perhaps rearrangement) may occur after that of the β -chain, analogous to light-chain immunoglobulin expression following that of the heavy chain in B cells³³.

How many T-cell receptor gene families?

Regardless of whether TT11 or HDS4, or neither, encodes the T-cell receptor α -chain, there is already an overabundance by

at least one of immunoglobulin-like T cell-specific genes which rearrange in some types of T-cells. The reason for this cannot be explained based on the present biochemical evidence, and may involve variations of a two-receptor model of T-cell recognition⁴⁹. Although the strict 'two-receptor' model for MHC-antigen co-recognition has been excluded by several convincing experiments^{44,50}, there exists the possibility that in at least some circumstances an alternative heterodimer might be able to form for some purpose related to this event.

We thank Dr J. Casey for the gift of random primer DNA, Dr C. Benoist for the thymocyte and spleen cDNA libraries, Dr C. Bona and colleagues for the FSN-1 and FSN-13 DNAs, Dr E. Ruley for 3T3 mRNA, Dr K. Arai and colleagues for the HDS4 probe, Dr J. Rothbard for the molecular weight calculation and G. Samuel and K. Redman for secretarial assistance. D.M.B. was supported in part by a fellowship from the Lupus Foundation, T.L. is a fellow of the Damon Runyon-Walter Winchell Cancer Fund and D.C. was a fellow of the Leukemia Society of America. This work was supported by grants to M.M.D. from NIH and the American Cancer Society.

Note added in proof: Amino acid sequencing of a human T-cell receptor α -chain (N. Jones and J. L. Strominger, personal communication) suggests that TT11 is the murine homologue. Also, most of the 8-10 bands seen on Southern blot (Fig. 3) with a TT11 probe hybridize to the variable portion of the molecule. This contrasts with results obtained with V_β probes²⁴ and suggests that V_α may be immunoglobulin-like.

Received 9 September; accepted 1 October 1984.

- Fathman, C. G. & Frelinger, J. G. *A. Rev. Immun.* **1**, 633-655 (1983).
- Zinkernagel, R. M. & Doherty, T. C. *Nature* **248**, 701-702 (1974).
- Shearer, G. M. *Eur. J. Immun.* **4**, 527-533 (1974).
- Bevan, M. J. *J. exp. Med.* **142**, 1349-1364 (1975).
- Benaceraf, B. & Germain, R. N. *Immun. Rev.* **38**, 70-119 (1978).
- Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882-896 (1978).
- Bevan, M. J. & Fink, P. J. *Immun. Rev.* **42**, 3-28 (1978).
- Infante, A. J., Infante, P. D., Gillis, S. & Fathman, C. G. *J. exp. Med.* **155**, 1100-1107 (1982).
- Allison, J. P., MacIntyre, B. W. & Block, D. J. *Immun.* **129**, 2293-2300 (1982).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Samelson, L. E., Germain, R. N. & Schwartz, R. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6972-6976 (1983).
- Kaye, J., Porcelli, S., Tite, J., Jones, B. & Janeway, C. A. Jr *J. exp. Med.* **158**, 836-856 (1983).
- Acuto, O. *et al. Cell* **34**, 717-726 (1983).
- Kappler, J. *et al. Cell* **35**, 295-302 (1983).
- Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149-153 (1984).
- Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
- Yanagi, Y. *et al. Nature* **308**, 145-149 (1984).
- Acuto, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3851-3855 (1984).
- Chien, Y., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
- Siu, G. *et al. Cell* **37**, 393-401 (1984).
- Kavalier, J., Davis, M. M. & Chien, Y. *Nature* **310**, 421-423 (1984).
- Siu, G. *et al. Nature* **311**, 344-349 (1984).
- Patten, P., Yokota, T., Rothbard, J., Chien, Y., Arai, K. & Davis, M. M. *Nature* **312**, 40-46 (1984).
- Saito, H. *Nature* **309**, 757-762 (1984).
- Samelson, L. A. & Schwartz, R. H. *Immun. Rev.* (in the press).
- MacIntyre, B. W. & Allison, J. P. *Cell* **38**, 654-665 (1984).
- Kaye, J. & Janeway, C. A. Jr *J. Immun.* (in the press).
- Yoshikai, Y., Yanagi, Y., Suciu-Foca, N. & Mak, T. W. *Nature* **310**, 506-508 (1984).
- Roy, H. D., Bensussan, A., Acuto, O. & Reinherz, E. L. *J. exp. Med.* (in the press).
- Hedrick, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Koren, H., Handwerker, B. & Wunderlich, J. J. *Immun.* **114**, 894-896 (1975).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Nagase, F., Waters, S. J., Thorbecke, G. J. & Bona, C. A. *Eur. J. Immun.* (in the press).
- Maxam, A. & Gilbert, W. *Methods Enzym.* **65**, 499-560 (1980).
- Gascoigne, N. R. J., Chien, Y., Becker, D. M., Kavalier, J. & Davis, M. M. *Nature* **310**, 387-391 (1984).
- Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
- Chien, Y., Kavalier, J. & Davis, M. M. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. in *Sequences of Proteins of Immunological Interest* (US Dept of Health and Human Services, Washington, DC, 1983).
- Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15**, 1495-1509 (1978).
- Parnes, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2253-2257 (1981).
- Davis, M. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2194-2198 (1984).
- Orkin, S. H. & Swerdlow, P. S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2475-2479 (1977).
- Kappler, J., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198-1214 (1981).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1983).
- Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D. M.) (IRL, Oxford, 1984).
- Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **442**, 324-330 (1976).
- Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
- Williamson, A. R. *Nature* **283**, 527-532 (1980).
- Heber-Katz, E. *et al. J. exp. Med.* **155**, 1198-1214 (1981).

A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes

Haruo Saito, David M. Kranz, Yohtaroh Takagaki, Adrian C. Hayday, Herman N. Eisen & Susumu Tonegawa

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

In addition to the two previously identified genes rearranged and expressed in a cytotoxic T-lymphocyte clone, we have identified a third gene that is also rearranged and expressed in the same clone. This new gene shows clonal diversity, codes for a polypeptide chain that contains immunoglobulin-like variable and constant domains, carries potential N-glycosylation sites and is a particularly attractive candidate for the gene that encodes the α -subunit of the heterodimeric antigen receptor of this T-cell clone.

IN spite of its central roles in both regulatory and effector phases of the vertebrate cellular immune system, the chemical nature of the T-cell antigen receptor has been elusive¹. A particularly intriguing question about this receptor concerns the molecular basis for its dual recognition of antigens in conjunction with self major histocompatibility complex (MHC) products²⁻⁴. Unlike immunoglobulins, which recognize and bind free antigens, the T-cell antigen receptors characteristically recognize cell-bound antigens in the specific molecular context of self-MHC molecules. Recently, several groups have succeeded in making antisera and monoclonal antibodies that are specific for individual T-cell clones⁵⁻⁷ and which presumably recognize clone-specific determinants on the T-cell antigen receptors. The receptor thus characterized is a heterodimeric glycoprotein of apparent relative molecular mass (M_r) ~90,000, consisting of an α -subunit (M_r 42,000-44,000) and a β -subunit (M_r 42,000-44,000) held together by an inter-chain disulphide bond(s). Peptide fingerprint analyses of α - and β -chains from several T-cell lines suggest that both subunits are composed of variable (V) and constant (C) regions⁸⁻¹⁰.

During the past year, three groups have isolated cDNA clones [YT35 (ref. 11), TM86 (ref. 12), pHDS11 (ref. 13) and pHDS4/203 (ref. 13)] that probably correspond to T-cell receptor genes. As YT35, TM86 and pHDS11 from a human T-cell tumour, a mouse T helper hybridoma and a mouse cytotoxic T lymphocyte (CTL) clone, respectively, are all extensively homologous to each other at the 3' half, it was concluded that they correspond to the same gene. Recently, partial amino acid sequences determined by direct analysis of the β -chain of a human T-cell receptor showed that YT35, TM86 and pHDS11 represent the gene for this subunit¹⁴. Although there is no extensive nucleotide homology between the fourth clone, pHDS4/203, and the other three cDNA clones, all share several common properties. (1) They are expressed in T cells but not in B cells. (2) The corresponding genes are rearranged in T cells and not in B cells. (3) In the encoded proteins there are distinctive regions that, proceeding from amino- to carboxy-terminus, correspond to a signal peptide, two immunoglobulin-like domains, a transmembrane peptide rich in hydrophobic amino acids and a short cytoplasmic peptide. (4) Their deduced amino acid sequences have low but significant homology to those of immunoglobulin chains. (5) Each gene is composed of separate V, J and C gene segments¹⁵⁻¹⁸ (in the cases of YT35 and TM86, the corresponding genes have also been shown to have D segments)^{19,20}. (6) Their V, (D) and J segments have characteristic signal sequences for gene rearrangement (heptamer and nonamer separated by either 12 or 23 base pairs (bp); refs 15, 16, 18-20 and A. Hayday *et al.*, unpublished data).

The cDNA clone pHDS4/203 therefore satisfied many of the

properties expected for the α -chain of the T-cell antigen receptor. However, this clone lacks sequences that correspond to sites for N-linked glycosylation and it has recently been shown that both the α -chain and the β -chain of another type of T cell, helper cells, are N-glycosylated (J. P. Allison, personal communication). Although it was possible that the α -chain of the CTL clone from which pHDS4/203 was isolated is not N-glycosylated, it also seemed possible that there is still another gene (or genes) whose properties are similarly consistent with those expected for the α -subunit of the T-cell receptor. Accord-

Table 1 T cell-specific clones from a subtracted CTL cDNA library

Group	Size of mRNA (kb)	No. of clones	Rearrangement	Representative clone
T (thyl)	2.0	19	No	pHDS1
B	0.8	19	No	
C(β)	1.4	18	Yes	pHDS11
F	1.7	13	No	
E	0.8	12	No	
D	1.7	9	Yes	pHDS58
K	1.3/1.1	9	No	
G	1.2	8	No	
A	1.5	4	Yes	pHDS4/203
P	2.3	2	No	
Q	n.d.	2	No	
I	2.2	1	No	
J	1.0	1	No	
L	1.5	1	No	
N	4.4	1	No	
R	0.8	1	Maybe	pHDS86
Total		120		

Two hybridization probes were used to screen the cDNA library. The first was the 2C cDNA prepared from the poly(A)⁺ RNA of membrane-bound polysomes by subtraction with poly(A)⁺ RNA from a B-cell lymphoma, CH-1. The second was the cDNA prepared from the total poly(A)⁺ RNA from another B-cell lymphoma A20-2J. After screening about 20,000 independent transformants, we have identified a total of 140 putative T cell-specific cDNA clones and grouped these clones on the basis of the corresponding genes, using the following procedures. First, plasmid DNA preparations from each of 20 randomly chosen cDNA clones were used as hybridization probes in RNA blotting analysis of 2C poly(A)⁺ RNA and DNA blotting analysis of 2C and embryo DNA. The hybridization patterns allowed assignment of these clones to 12 groups. Second, mouse DNA inserts were dissected from a representative cDNA clone of each group and used as probes for colony hybridization of all of the cDNA clones; the results allowed assignment of 105 clones among the 12 groups. The above procedures were repeated for some of the remaining unassigned clones.

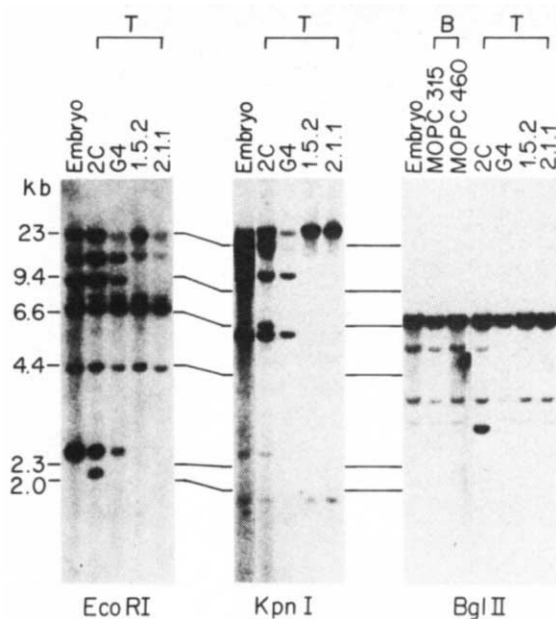


Fig. 1 Southern blot analyses²² of DNA from BALB/c embryo, myelomas MOPC315 and MOPC460 (both BALB/c-derived) and CTL lines 2C, G4, 1.5.2 and 2.1.1 (all BALB.B-derived²¹). DNA was digested with the indicated restriction enzymes, electrophoresed through 0.8% agarose gels, blotted onto nitrocellulose and hybridized with the ³²P-labelled, nick-translated insert from clone pHDS58. Hybridization was carried out at 42 °C in 50% formamide and 5×SSC. The filters were washed at 65 °C in 0.2×SSC. Separate experiments showed that BALB/c and BALB.B embryo patterns are indistinguishable. Numbers on the left-hand side are *M_r* markers.

ingly, we have analysed further the cDNA library from CTL clone 2C and report here the isolation of a third class of cDNA clone representing a gene whose protein product is a better candidate for the α -subunit of the T-cell $\alpha\beta$ heterodimer receptor than is pHDS4/203: the new gene is rearranged and expressed specifically in T cells, its product has the sequences expected of an integral membrane protein and it carries potential N-glycosylation sites.

Classification of cDNA clones

We reported previously¹³ construction of a 'subtracted' cDNA library from an alloreactive CTL clone (2C) specific for the product of a gene at the *D* end of the *H-2* complex²¹. We subjected this library to differential screening using two hybridization probes (Table 1). From these experiments, it was established that 120 of 140 putative T cell-specific cDNA clones originally identified, belonging to 16 different groups, are expressed specifically in CTL clone 2C and not in B cells. Table 1 lists the number of cDNA clones in each group and the size of poly(A)⁺ RNA they detected by Northern hybridization. Of the remaining 20 cDNA clones, 12 belonged to two gene groups whose expression in B cells was substantially lower than in T cells but still definitely detectable. The other eight cDNA clones contained relatively short inserts (less than 250 bp) and have not yet been characterized.

Among the 16 T cell-specific cDNA groups, we identified four whose genes seemed to be rearranged in CTL clones. Two (pHDS11 and pHDS4/203) of the four have been described previously¹³. The characteristics of the third group, represented by pHDS58, are the main subject of the present report. The fourth cDNA, represented by one member, pHDS86, hybridized to an extra DNA fragment when *Pvu*II digests of clone 2C DNA and embryo DNA were compared; however, nine other restric-

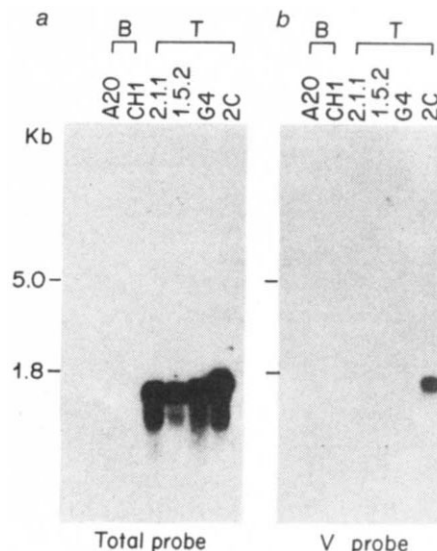


Fig. 2 RNA blot analyses²³ of poly(A)⁺ RNA from various B- and T-cell lines. Probes: *a*, total insert of pHDS58; *b*, *Hpa*II fragment containing 270 bp at the end of the pHDS58 insert (V-region probe). RNA was extracted from B-cell lymphomas A20-2J and CH1 and alloreactive (H-2^b anti-H-2^d) cytotoxic T-lymphocyte clones 2.1.1, 1.5.2, G4 and 2C. Approximately 1.5 µg of poly(A)⁺ RNA was denatured with glyoxal and electrophoresed through a 1% agarose gel in 10 mM sodium phosphate buffer (pH 6.5). RNA was transferred to a nitrocellulose membrane and hybridized to ³²P-labelled nick-translated probe DNAs. The conditions of hybridization and washing are given in Fig. 1 legend. The filter used for hybridization with the first probe was washed by boiling at 100 °C for 5 min in H₂O and re-used for hybridization with the second probe.

tion enzymes failed to show any evidence of rearrangement (data not shown).

Gene rearrangement in T cells

Figure 1 shows the results of Southern blot analyses²² of DNA from BALB/c embryos, four CTL clones of different specificities and two myelomas that have been digested with *Eco*RI, *Kpn*I or *Bgl*II and hybridized with the insert of pHDS58. With each of the three enzymes, clone 2C DNA showed an extra fragment that was not detected in the digest of embryo DNA. DNA from the other CTL clones gave hybridization patterns which lacked several fragments detected in embryo DNA. In contrast, the hybridization patterns of DNA from the myelomas and embryos were indistinguishable. These results strongly suggest that the gene or genes corresponding to clone pHDS58 are rearranged in cytotoxic T cells but not in myelomas.

The expression of the pHDS58 genes in CTL clones and B lymphomas was examined by Northern hybridization analyses²³ of poly(A)⁺ RNA. When the whole insert of pHDS58 was used as the hybridization probe, no hybridization was detected in the two B-cell lymphomas, but poly(A)⁺ RNAs of distinct sizes were detected in all four CTL clones (Fig. 2*a*). The size of the major RNA species in 2C (1.7 kilobases, kb) was slightly greater than that in the other three CTL clones (1.6 kb). In addition, minor RNA species of about 1.2 kb were detected in all four CTL clones. When a DNA fragment containing 270 bp from the 5' end of the pHDS58 insert was used as the hybridization probe, only the 1.7-kb RNA species of 2C was detected and the other three CTL clones and the B-cell lymphomas exhibited no hybridization (Fig. 2*b*; see Fig. 3 for the region of the cDNA sequence covered by the probe). These results suggest that the 1.6-1.7-kb RNA is composed of a common 3' sequence (C region) and a CTL clone-specific 5' sequence (V region), as has been seen with immunoglobulin and the T-cell receptor β -chain.

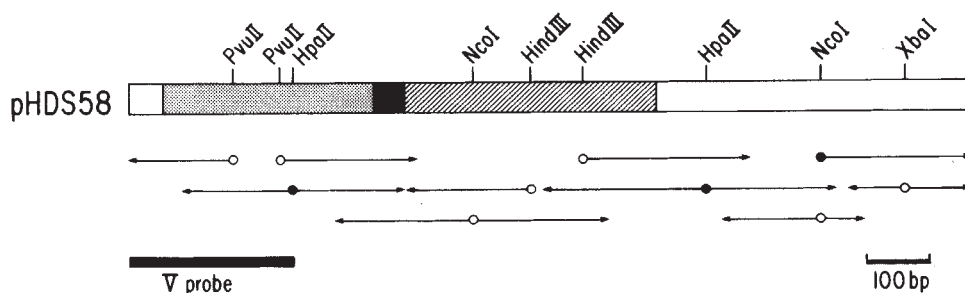


Fig. 3 Restriction map of the insert of cDNA clone pHDS58. The map was constructed by a combination of single and double restriction enzyme digestions of the plasmid DNA. The V, J, C and 5'- or 3'-untranslated regions are shown by shading, black, hatching and white, respectively. Also shown is the sequencing strategy used to determine the nucleotide sequence shown in Fig. 4. The arrows indicate the directions and the extent of sequence determination. The open and closed circles indicate that the ends of the DNA fragments labelled are 5' and 3', respectively.

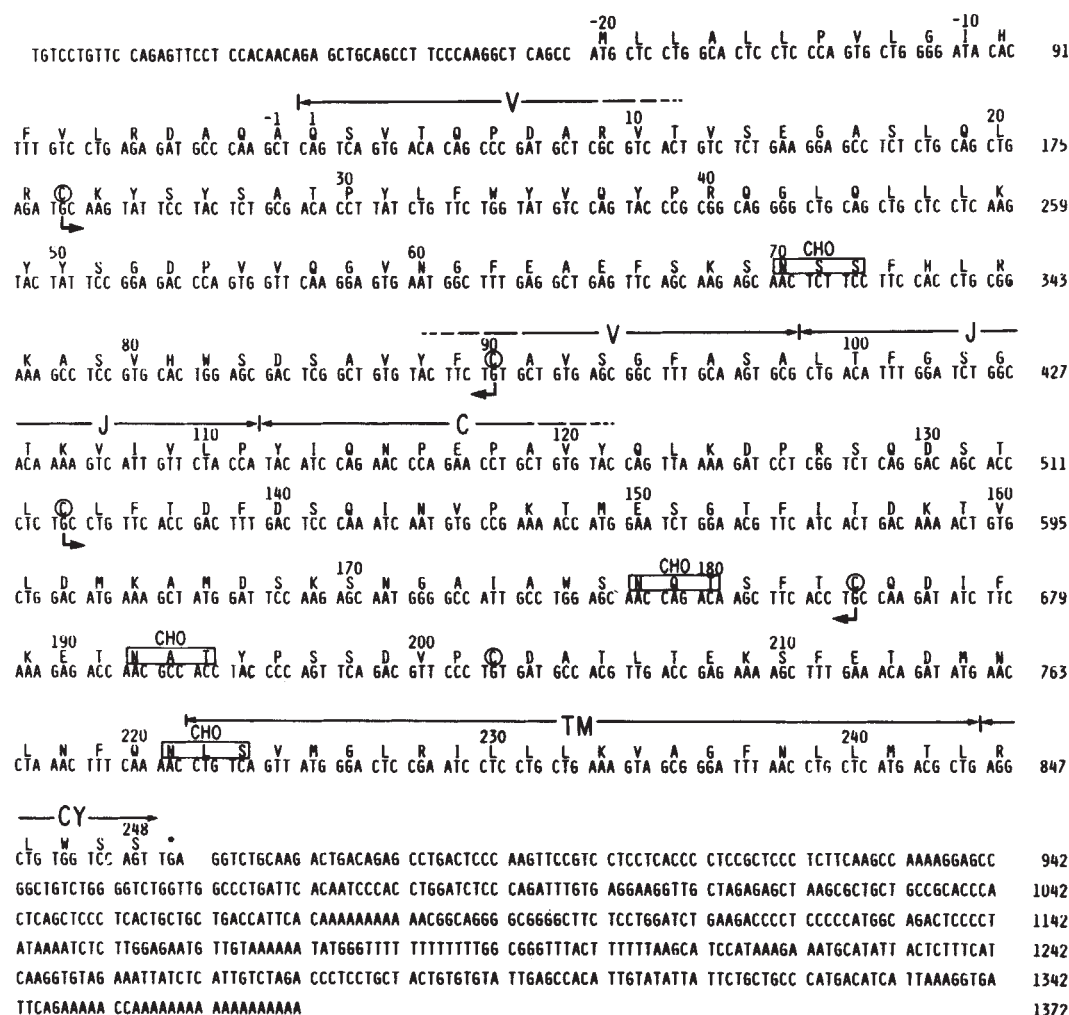


Fig. 4 The nucleotide and predicted amino acid sequences of the cDNA clone pHDS58. The nucleotide sequence was determined by the method of Maxam and Gilbert²⁴ according to the strategy shown in Fig. 3. Numbers given above the amino acid sequence designate amino acid residue positions. The numbers on the right show nucleotide positions. The V, J, C, TM (transmembrane) and CY (cytoplasmic) regions are indicated by horizontal arrows although exact boundaries are ambiguous. Cysteine residues thought to be involved in intradomain or inter-chain disulphide linkages are encircled. The potential N-glycosylation sites (N-X-S or N-X-T) are also indicated.

Nucleotide sequence

The restriction map and the sequencing strategy of the pHDS58 insert are shown in Fig. 3. The entire nucleotide sequence of 1,372 bp determined by the method of Maxam and Gilbert²⁴ is shown in Fig. 4. The longest open reading frame begins with the Met codon at nucleotide positions 56–58, extends over a stretch of 804 bp, and ends at nucleotide position 859. The amino

acid sequence corresponding to this reading frame is also shown in Fig. 4. For reasons given below, the numbering of codons or amino acid residues starts with the triplet CAG (Gln) at nucleotide positions 116–118. The amino acid sequence immediately following the Met and extending to the Gln is highly hydrophobic and reminiscent of a signal peptide²⁵. After the Gln, the sequence is significantly homologous to those of the T-cell

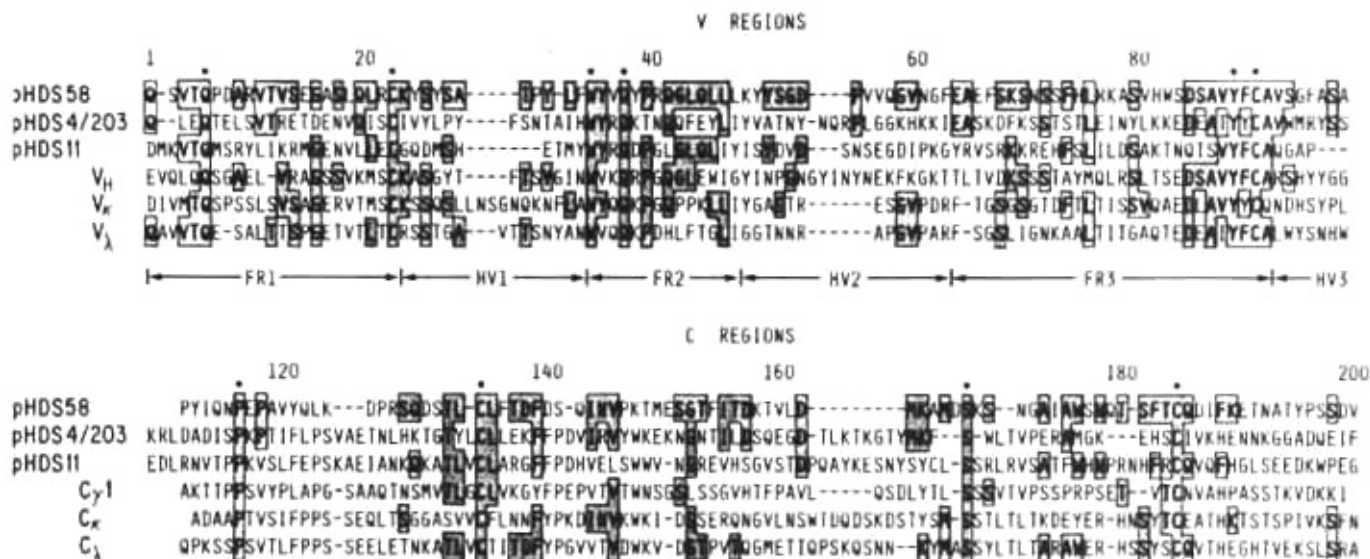


Fig. 5 Comparisons of the V and C regions of the predicted amino acid sequence of pHDS58 and the V and C regions, respectively, of the predicted or determined amino acid sequences of five other polypeptide chains: a chain encoded by pHDS4/203, the β -chain encoded by pHDS11 (ref. 13), 93G7 γ 1 immunoglobulin heavy chain, MOPC603 κ light chain, and MOPC104E λ 1 light chain²⁶. Those residues common between the pHDS58 protein and any one of the other five chains are boxed. *. The residues common among all six chains. Approximate boundaries of framework and hypervariable regions as they appear in immunoglobulin V regions are indicated by horizontal arrows. The amino acid positions refer to those of pHDS58 protein.

receptor β -chain deduced from pHDS11 and the polypeptide chain encoded by pHDS4/203, as well as to the sequences of an immunoglobulin heavy chain (93G7), an immunoglobulin κ light chain (MOPC603) and an immunoglobulin λ 1 light chain (MOPC104E)²⁶.

The homology is shown in Figs 5 and 6, where the pHDS58 amino acid sequence is compared with the V (or V+D), J and C region sequences of the other five polypeptide chains. Sequence homology is evident in all three regions, but is most striking in the J region, where it amounts to 38–62%. In the V regions, the pHDS58 sequence is related to the other five sequences by 22–29% homology. The six amino acid residues that are conserved in this region among pHDS11, pHDS4/203 and the three types of immunoglobulin chains (Gln, Cys, Trp, Gln, Tyr and Cys at positions 5, 22, 34, 37, 88 and 90, respectively) are also shared by pHDS58. The sequence homology is least in the C regions, where it ranges over 12–20%. Nevertheless, the relatedness is evident around the two Cys residues that form intradomain disulphide bonds in immunoglobulin chains. Throughout the V, J and C regions, the pHDS58 sequence, like the pHDS11 and pHDS4/203 sequences, is more homologous to the light chains (25% and 27% for κ and λ , respectively) than to the heavy chain (23%).

Beyond the C region, the pHDS58 polypeptide chain exhibits no obvious sequence homology to the corresponding regions of the β -chain (pHDS11) or the chain encoded by pHDS4/203. Nevertheless, the three polypeptide chains are organized in a very similar fashion in these regions. Thus, as in the pHDS11 and pHDS4/203 chains, the C region of the pHDS58 chain is followed by a peptide carrying an extra cysteine, then by a stretch of about 20 hydrophobic residues that could correspond to a transmembrane peptide and finally by a short hydrophilic C-terminal peptide that presumably extrudes into the intracytoplasmic space.

In Fig. 4 these regions of pHDS58 are indicated by horizontal arrows. The exact boundaries of a few adjacent regions are somewhat uncertain. For instance, it is not possible to determine unambiguously the N-terminus of the processed protein from the nucleotide sequence of the cDNA alone. However, the Gln at nucleotide positions 116–118 is the best candidate for the following reasons. First, the proposed assignment places the first Cys residue at position 22, while the corresponding Cys

residues of most immunoglobulin chains are at position 22 or 23 (ref. 26). Second, most immunoglobulin chains carry a signal peptide 19–22 residues long²⁶; the present assignment makes the proposed signal peptide 20 residues long. Finally, the pHDS58 chain is an excellent candidate for the α -subunit of the T-cell receptor, whose N-terminal residue has recently been shown to be blocked (S. Schlossman, personal communication), probably by a cyclized glutamine residue.

The proposed processed polypeptide chain is 248 amino acids long and its calculated relative molecular mass is ~28,000. The chain contains 23 negatively charged (15 Asp and 8 Glu) and 19 positively charged residues (6 Arg and 13 Lys), corresponding to an isoelectric point near neutrality in the absence of post-translation modifications. As shown in Fig. 4, it has four potential sites for N-glycosylation.

Conclusions

The pHDS58 gene and its product share several characteristics with the two previously identified T-cell genes, pHDS11 and pHDS4/203, and their products. (1) The gene is specifically expressed in various CTL clones but not in B lymphomas. (2) The gene is rearranged in CTL clones but not in myelomas and the rearrangement pattern varies with the CTL clone. (3) The corresponding poly(A)⁺ RNA is composed of 5'-variable and 3'-constant regions. (4) The primary sequence suggests that the encoded protein is composed of a signal peptide, two immunoglobulin-like domains, each with one disulphide loop,

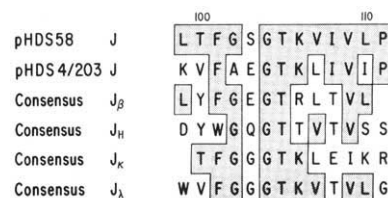


Fig. 6 Comparison of the J region amino acid sequences of pHDS58 and pHDS4/203 (ref. 13) and consensus sequences of the T-cell antigen receptor β -chain¹⁶ and immunoglobulin J_H , J_κ and J_λ (ref. 26). Those residues common between pHDS58 and any one of the other five J segments are boxed and shaded.

a transmembrane peptide and a cytoplasmic peptide. (5) The two domains are homologous to the V and C domains of immunoglobulin chains, particularly those of λ light chains. (6) Besides having two cysteine residues in each domain, the encoded protein has a fifth cysteine residue in the region between the C domain and the transmembrane peptide, that is, at a position where the chain might be disulphide bonded to another chain¹³.

The counterparts of the pHDS11 gene in a mouse T-helper hybridoma and a human T-cell tumour have been identified previously by Hedrick *et al.*¹² and Yanagi *et al.*¹¹ and shown to code for the β -subunit of the heterodimeric T-cell receptor¹⁴. As the nucleotide sequences of pHDS11 and the gene identified by Hedrick *et al.*¹² are virtually identical in the C region, we concluded that pHDS11 codes for the β -subunit of the CTL receptor¹³.

Because pHDS4/203, the second type of rearranged gene discovered in the CTL clone 2C, had many characteristics expected for a T-cell receptor gene, it was originally thought to code for the other (α) subunit of the CTL receptor¹³. Subsequently, endoglycosidase F digestion of the $\alpha\beta$ heterodimers of mouse helper T cells and a human T-cell tumour suggested that both subunits of these T-cell receptors are N-glycosylated (J. P. Allison, personal communication; C. A. Janeway, personal communication), but the pHDS4/203 gene, unlike the pHDS11 gene, has no typical N-glycosylation sites (namely, the tripeptide

Asn-X-Ser/Thr)¹³. Although the difference in the source of T cells (that is, CTL versus helper and mouse versus human) could have explained the apparent discrepancy, our recent studies suggest that the α -subunit of CTL clone 2C is also N-glycosylated (unpublished data). It is therefore of particular interest that the newly discovered pHDS58 gene reported here not only possesses all of the properties expected for a T-cell receptor gene but also carries four potential N-glycosylation sites. Thus, while there is no direct evidence, the pHDS58 gene must be considered a better candidate than pHDS4/203 for the α -subunit.

If the pHDS58 gene encodes the α -subunit and pHDS11 the β -subunit of 2C clone $\alpha\beta$ heterodimeric receptor, what is the function of the pHDS4/203 gene? As this gene exhibits clonal diversity, in common with the other two genes, it also is likely to be involved directly in determining T-cell specificity. If so, this raises the possibility of a second receptor. The idea of a second receptor has been previously discussed at length²⁷⁻³² but is at present less favoured than the one-receptor hypothesis³³⁻³⁶. The role of the third gene should be clarified by analysis of its product, currently underway.

We thank John McMaster and Mark Pichler for technical assistance, Eleanor Basel for secretarial assistance and Sifri Novotny for discussions. The work was supported by grants CA-28900-04 (S.T.), CA-28900 (H.N.E.) and CA-14051 (a core grant to S. E. Luria). D.M.K. is supported by the Arthritis Foundation.

Received 4 October; accepted 11 October 1984.

- Jenselius, J. C. & Williams, A. F. *Nature* **300**, 583-588 (1982).
- Kindred, B. & Shreffler, D. C. *J. Immun.* **109**, 940-943 (1972).
- Katz, D. H., Hamaoka, T., Dorf, E. D., Maurer, P. H. & Benacerraf, B. *J. exp. Med.* **138**, 734-739 (1973).
- Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701-702 (1974).
- Allison, J. P., McIntyre, B. W. & Bloch, D. J. *Immun.* **129**, 2293-2300 (1982).
- Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Acuto, O. *et al. Cell* **34**, 717-726 (1983).
- McIntyre, B. W. & Allison, J. P. *Cell* **34**, 739-746 (1983).
- Kappler, J. *et al. Cell* **35**, 295-302 (1983).
- Yanagi, Y. *et al. Nature* **308**, 145-149 (1984).
- Hedrick, S. M., Nielsen, E. A., Kavaler, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
- Saito, H. *et al. Nature* **309**, 757-762 (1984).
- Acuto, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3851-3855 (1984).
- Chien, Y., Gascoigne, N. R. J., Kavaler, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
- Gascoigne, N. R. J., Chien, Y., Becker, D. M., Kavaler, J. & Davis, M. M. *Nature* **310**, 387-391 (1984).

- Siu, G. *et al. Cell* **37**, 393-401 (1984).
- Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
- Kavaler, J., Davis, M. M. & Chien, Y. *Nature* **310**, 421-423 (1984).
- Siu, G. *et al. Nature* **311**, 344-350 (1984).
- Kranz, D. M., Sherman, D. H., Sitkovsky, M. V., Pasternack, M. S. & Eisen, H. N. *Proc. natn. Acad. Sci. U.S.A.* **81**, 573-577 (1984).
- Southern, E. J. *molec. Biol.* **98**, 503-517 (1975).
- Thomas, P. S. *Meth. Enzym.* **100**, 255-266 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835-851 (1975).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Proteins of Immunological Interest* (NIH, Bethesda, 1983).
- Von Boehmer, H., Haas, W. & Jerne, N. K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2439-2442 (1978).
- Blanden, R. V. & Ada, G. L. *Scand. J. Immun.* **7**, 181-190 (1978).
- Langman, R. E. *Rev. Physiol. Biochem. Pharmac.* **81**, 1-37 (1978).
- Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882-896 (1978).
- Cohn, M. & Epstein, R. E. *Cell. Immun.* **39**, 125-153 (1978).
- Janeway, C. A. Jr, Wigzell, H. & Binz, H. *Scand. J. Immun.* **5**, 993-1001 (1976).
- Matzinger, P. *Nature* **292**, 497-501 (1981).
- Hunig, T. R. & Bevan, M. J. *J. exp. Med.* **155**, 111-125 (1982).
- Hedrick, S. M. *et al. Cell* **30**, 141-152 (1982).
- Kappler, J. W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198-1214 (1981).

Structure, expression and divergence of T-cell receptor β -chain variable regions

Phillip Patten*, Takashi Yokota†, Jonathan Rothbard‡, Yueh-hsiu Chien‡, Ken-ichi Arai† & Mark M. Davis*

† Department of Medical Microbiology and * Department of Biological Sciences, Stanford University School of Medicine, Stanford, California 94305, USA

‡ DNAX Research Institute for Molecular and Cellular Biology, Palo Alto, California 94306, USA

Analysis of three new T-cell receptor β -chain variable regions together with those in the literature indicates that they have both remarkable similarities and differences with those of immunoglobulin. Less than 10 V regions appear to predominate in the thymus. V_{β} sequences are much more heterogeneous at the amino acid level than are immunoglobulin V regions and they appear to diverge between species much more quickly, apparently the result of additional hypervariable regions. Three of these putative new hypervariable regions lie outside of the classical immunoglobulin binding site, an indication that important interactions may be occurring in these regions with polymorphic MHC determinants.

RECOGNITION of foreign entities by T cells appears to involve molecules and mechanisms strikingly similar to those involving immunoglobulins. Analysis of T-cell receptor β -chain cDNA clones and the genes encoding them has revealed V, D, J and

C regions similar in size and sequence to the corresponding immunoglobulin elements¹⁻¹⁰. The V, D and J genomic elements are rearranged during T-cell differentiation to form a single exon, as is the case for immunoglobulins, and this rearrangement

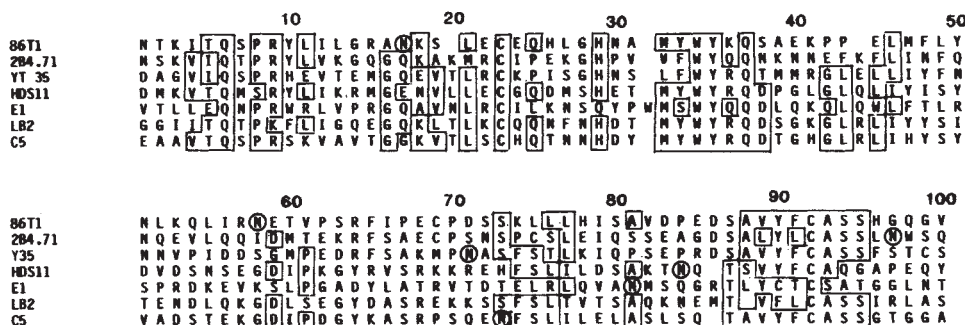


Fig. 2 Amino acid homologies between the seven β -chain variable-region sequences available. Homologies between four or more sequences are boxed. Both V- and D-region sequences are shown. Potential N-linked glycosylation sites are circled.

cDNA clones with complete V_{β} coding regions were isolated from the E1, C5 and LB2 libraries and their nucleotide sequences were determined by the method of Maxam and Gilbert²² according to the strategy in Fig. 1. All of these transcripts derive from the 3' most $J-C$ locus ($J_{T\beta}-C_{T\beta}$)⁶, as do the sequences of all four of the other full-length cDNAs derived from functional lines, hybrids or lymphomas^{4,23,24}. Although cDNA clones in the spleen or thymus appear to rearrange equally well to either of the two loci, at least in the context of a specific V region²⁴, few functional T-cell lines appear to use the 5' most locus (G. K. Sim, personal communication, N. R. J. Gascoigne and J. Elliot, personal communication). As the two C regions are almost identical⁶, the strong bias of such cells for the 3' most locus may reflect a selection for J regions from that cluster and may be an important clue in determining the functional significance of a tandem $J-C$ arrangement of the B -chain locus.

The data presented in Fig. 1 also show that E1 and C5, though reactive with structurally related antigens, share very little sequence homology. This lack of homology may reflect the effect of carrier, as C5 is conjugate-specific for DNP-ovalbumin whereas E1 is hapten-specific for TNP. Alternatively, it could reflect the effect of different restricting alleles (E1, I-A^d; C5, I-A^b) on antigen binding. Another explanation is that the other chain of the receptor (α) binds antigen in these clones. In addition to the lack of homology between the β -chain variable regions of E1 and C5, there is no homology in any portion of these with the antigen-binding portions of known TNP- or DNP-binding immunoglobulins²⁵—a further indication of the very different requirements of T-cell and B-cell recognition and an indication of why receptor isolation schemes involving T- and B-cell lines of matched antigen specificity were unsuccessful²⁶⁻²⁸.

β -Chain V regions in thymus

To assess the relative contributions of various β -chain V regions to the T-cell receptor repertoire, V region-specific fragments were prepared from each of the three cDNA clones described

above; from 2B4.71, a cytochrome c -specific T_h hybrid⁴; and from 86T1 (ref. 2), isolated from a thymocyte cDNA library (see legend to Table 1). These five distinct V -region fragments were labelled and used to screen 250,000 λ gt10 (ref. 29) cDNA clones from a BALB/c strain thymocyte poly(A)⁺ RNA-derived cDNA library (200,000 independent clones). Table 1 shows the percentage of clones in this library that were positive for both the β -chain constant region and these V regions. This analysis indicates that four of the five V regions fall into a uniformly frequent class of transcripts, each representing 1–3% of the C -region hybridizing clones. Since four of these five genes hybridize to single bands, and the fifth to only two bands in genomic Southern blots (see below), these frequencies must largely reflect the expression levels of the loci from which they were derived. The actual frequency of these V_{β} sequences is even greater since, of 35 C_{β} -positive clones analysed from this library, only six extended sufficiently far 5' to have hybridizable V regions (data not shown), presumably due to frequent $D-J$ joining events^{7,9} and a 3'-end bias in cloning. The corrected frequencies are indicated in Table 1. Therefore, these four V regions together comprise 8.3% of the total β -chain message, and an even larger fraction (~40%) of functional transcripts. Thus, a relatively small number of V regions, perhaps less than ten, predominate in the thymus. The notion of a small V_{β} repertoire is supported by the fact that among the six murine V_{β} s reported here, two have been isolated independently from other functional T cells (86T1²⁴ and LB2, S. Hedrick, personal communication).

In contrast, the 2B4 V region was not detected among 250,000 clones surveyed, and thus must be expressed at a frequency at least 20 times lower. This is consistent with the observation that cytochrome c is a relatively weak antigen and reactive cells are only observed in secondary responses (R. Schwartz, personal communication). There are many such weakly reactive antigens^{30,31} and they may represent a second group of infrequently occurring V regions of unknown repertoire size. The regulatory mechanisms generating such uniformity in the frequency of the abundant class as opposed to the 2B4.71 V region are unknown. This selective expression contrasts with the much more diverse expression of immunoglobulin V regions in B cells, although a single heavy-chain variable region (V_H) seems to be expressed very frequently in early B cells (F. Alt, personal communication).

Additional hypervariable regions

With the sequences presented here, there are now seven β -chain V regions available^{2-4,23} for comparison with respect to framework and hypervariable regions. Homologies between the seven sequences are shown in Fig. 2 and the relevant numerical

Table 1 Variable region usage in a BALB/c thymocyte cDNA library

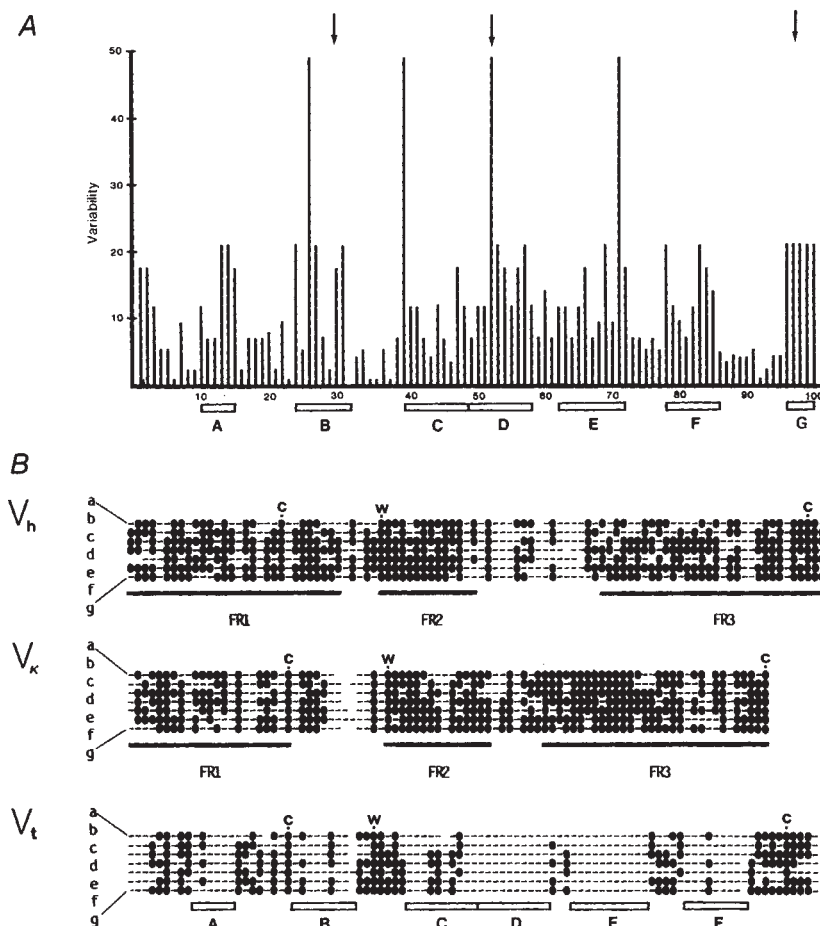
Variable region	Frequency (%)	Corrected frequency
C5	3.2	9.3%
E1	2.2	12.8%
LB2	0.9	5.2%
86T1	2.0	11.6%
2B4.71	<0.05	—
Total	8.3	38.9%

The following V region and C_{β} -specific fragments were purified on polyacrylamide gels and nick-translated by standard procedures⁴⁷: 86T1, *EcoRI*-*Bam*HI; 2B4.71, *EcoRI*-*Rsa*I; C5, *Pst*I-*Pvu*II; E1, *Pst*I-*Pst*I; LB2, *Bst*EII-*Xho*I; C_{β} 86T5², *EcoRI*-*EcoRI*. There were 1,650 C_{β} -positive clones. The percentage of these that were positive for a given V region is indicated. The corrected frequency is the raw percentage divided by the fraction of C_{β} -positive cDNA clones bearing V regions. The corrected value for C5 was divided by two as it hybridizes to two bands in genomic Southern blots and hence may cross-hybridize with another expressed V region.

Table 2 Variable-region homologies (%)

	2B4	YT35	HDS11	E1	LB2	C5
86T1	36	33	28	18	30	28
2B4		44	18	18	23	21
YT35			26	19	26	35
HDS11				18	44	51
E1					18	20
LB2						48

Fig. 3 Hypervariable and framework regions. **A**, Variability plot of the T-cell receptor β -chain V-region sequences shown in Fig. 2, using the convention of Wu and Kabat¹³. The positions of the three hypervariable peaks of immunoglobulin light chains²⁵ are indicated by arrows. Regions of hypervariability are marked by open bars. **B**, **D** and **G** correspond to the classical immunoglobulin hypervariable regions²⁵. Part **B** indicates the homologies between a variety of V_H , V_K and the seven V_β sequences. Homologies of four or better are indicated by a closed circle, less than four homologies by a dash, and spaces denote gaps inserted to maximize homology. Framework regions of the immunoglobulin sequences are indicated with solid bars. **A**, **C**, **E** and **F** are regions of V_β sequence that correspond to conserved V_K framework residues, but are significantly more variable in V_β . With the exception of the V_H region of *Caiman crocodylus*³³, sequences are from ref. 25. Immunoglobulin heavy chain (V_H): a, mouse group I MOPC 141CL; b, mouse group II MOPC 104E; c, mouse group III TEPC 601; d, dog MOO; e, rabbit K25; f, human VH26; g, caiman pC3. Immunoglobulin κ -chain (V_K): a, mouse 10K26-12A1; b, mouse S107B CL; c, dog GOM; d, human group I ROY; e, human group III, 1TI; f, rabbit K9 335; g, rabbit K25. V_T : a, 86T1; b, 2B4.71; c, YT35; d, HDS11; e, E1; f, LB2; g, C5.



comparisons in Table 2. Whereas the least related immunoglobulin V regions are 45% homologous in amino acid sequence and range up to 98% (ref. 25), the most similar V_β regions shown here are 51% homologous and range as low as 18% (Table 2). Although it has been suggested²³ that some of this variability is due to helper and cytotoxic T cells using different sets of V regions, this comparison shows that equally low homologies are observed between T_H -derived V_β 's, and that the cytotoxic T cell-derived V_β HDS11 has 51% amino acid homology with the T_H -derived V_β C5—the highest homology observed between any two V_β s in Table 2. Furthermore, the helper-derived V_β LB2 is nearly identical to nucleotide sequence (one change in 110 nucleotides examined) to a V_β derived from a functional cytotoxic T cell (S. Hedrick, personal communication).

These relatively low levels of homology among β -chain V regions are further analysed in the Wu-Kabat variability plot in Fig. 3A; this follows the formula (no. of different amino acids)/(frequency of the most common amino acid)¹³ and serves to highlight regions of maximum variability. Both heavy- and light-chain immunoglobulins exhibit three hypervariable peaks when plotted in this manner, two in the main body of the V region (CDR1 and CDR2) and a third corresponding to the V-J joint in light chains and the V-D-J joint in heavy chains²⁵ (indicated by arrows in Fig. 3A). In contrast, the V_β sequences examined in Fig. 3 exhibit seven distinct variability peaks, three of which correspond almost exactly to the canonical immunoglobulin κ light-chain hypervariable regions. Two of the peaks (A and F) correspond to regions in V_H and V_K that have been shown to be significantly more variable than classical framework regions^{25,32} but less variable than the classical hypervariable regions. The third and fifth peaks (C, E), however, do not have any precedent in immunoglobulins.

Complementary to these variability data are the homology comparisons shown in Fig. 3B; in this representation, four or more identical residues out of seven are displayed as solid circles

and fewer than four as dashes. In a control for the validity of conclusions based on a small number of sequences, a similar comparison is made of seven V_K and V_H sequences^{25,33}, chosen for maximum variability. In this analysis, the classical hypervariable (HV) and framework (FR) regions of immunoglobulins (indicated by solid bars under immunoglobulin sequences) are picked up remarkably well. By comparison, the β -chain V regions have stretches of homology flanking the conserved tryptophan and cysteines that are comparable with those in immunoglobulin FR regions. However, there are four new regions of sequence, corresponding to the four new peaks of variability in Fig. 3A, whose V_K equivalents are conserved FR residues, but which are much more variable in V_β (Fig. 3B, regions A, C, E, F). As discussed below, the sequence variability in these regions may have important consequences for antigen-MHC interactions.

Rapid divergence of β -chain V regions

There are two interpretations that are consistent with the remarkably low levels of homology between the V_β s reported here relative to immunoglobulin V regions. Cellular immunity could significantly predate humoral immunity and V_β s could evolve at a rate comparable to immunoglobulin V regions; the difference in homology between V_β s being then a result of having diverged from each other much earlier than immunoglobulin V regions. Alternatively, there could be significantly different selective pressures on V_β s as a consequence of the requirements of Ag-MHC interactions unique to the T-cell receptor. In the case of immunoglobulin, those portions of the V_H regions that interact with antigen (HV regions) acquire mutations at a much greater rate than do framework regions. Specifically, HV regions appear to diverge at the rate of unselected DNA, whereas FR regions evolve at a rate comparable to globin genes³⁴. These HV regions make up 22% of both V_H and V_K ²⁵, whereas the analysis in Fig. 3 shows that ~54% of V_β may be hypervariable. If these putative new HV regions evolve at a rate comparable to those

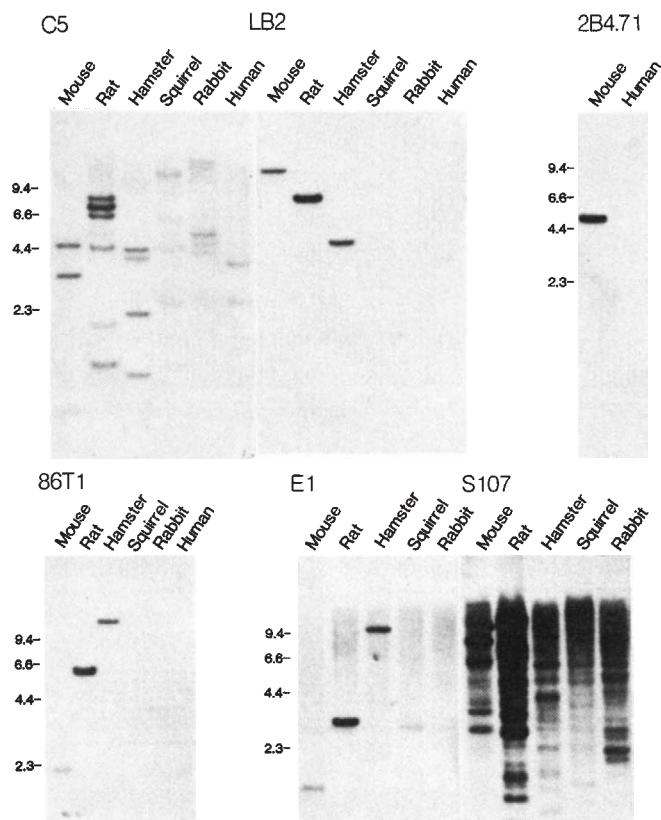


Fig. 4 Interspecies homologies of β -chain variable regions. *Eco*RI digests of mouse (BALB/c), rat, hamster (CHO), ground squirrel, rabbit and human DNAs were electrophoresed through 0.9% agarose and blotted onto nitrocellulose by standard procedures⁴⁷. V-region specific fragments from five of the β -chain clones and one from an immunoglobulin heavy chain (S107 V), were nick-translated and hybridized overnight at 55 °C in $6 \times$ SSC and washed at the same temperature, the final wash using $0.2 \times$ SSC for 30 min. Size markers are indicated in kilobases. Divergence time from mice (millions of years)^{36,37}: rat, 15 Myr; hamster, 25 Myr; ground squirrel, 60 Myr; rabbit, 70 Myr; human, 90 Myr.

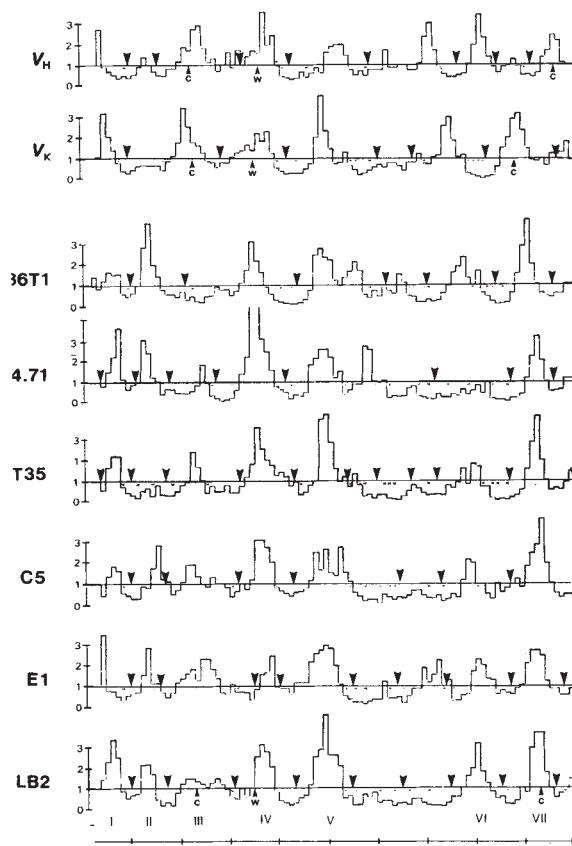


Fig. 5 Secondary structure predictions. The algorithm of Chou and Fasman⁴² was used to predict the locations of probable β -sheets and turns in the V_β sequences and in representative immunoglobulin sequences²⁵. P_B , plotted on the vertical axis, represents the β -sheet conformational parameter as defined⁴². Arrows are centred on regions of sequence where $P_T > P_B$ and $P_T > 0.75 \times 10^{-4}$. In regions where turns were predicted to be directly adjacent or to share amino acids, only the most probable turn is indicated by an arrow.

of immunoglobulin, the V_β gene as a whole should evolve much more rapidly than V_H or V_K genes.

To resolve between these interpretations, we used nucleic acid hybridization as an assay for the divergence of five murine V_β s and a representative V_H (S107)³⁵ from their homologues in other species. We reasoned that if a much larger fraction of V_β s is indeed hypervariable, hybridization of V region-specific probes with their homologues in successively more distantly related species should fall off much more rapidly for V_β than for V_H . Figure 4 shows a survey of six mammalian genomes, spanning 90-Myr of evolutionary time, probed on Southern blots with purified V_β probes and the S107 V region. Evidence that the rate of divergence of V_β s is in fact greater than that of V_H s is indicated by the fact that in general the V_β probes only hybridize well with closely related rodent genomes^{36,37}. S107 V_H , on the other hand, hybridizes well with squirrel (60 Myr divergence)³⁷ and rabbit (70 Myr divergence)³⁶ on comparable blots. Furthermore, S107 V_H and an independent V_H probe, μ 107 V (discussed in ref. 38), have been shown to hybridize well in more stringent conditions to much more distantly related genomes³⁸, including the primitive reptile, *Caiman crocodylus* (300 Myr divergence from mouse³⁶). Murine V_K probes have also been shown to cross-hybridize well on human genomic Southern blots³⁹. Taken together, these data indicate that most V_β s acquire mutations at a significantly greater rate than V_H s and V_K s¹⁵. While Southern blots do not provide the resolution of sequence data, this observation is consistent with the notion that a greater portion of V_β s are devoted to interaction with variable determinants and are thus selected to be hypervariable.

Consistent with an increased rate of divergence is the fact that four of the five V_β regions hybridize to single bands in the genome compared with between 4 and 30 bands that are observed with V_H ⁴⁰ and V_K ^{39,41} probes. That we see single hybridizing bands does not rule out the possibility that there are families of related V_β sequences; it suggests, rather, that members of a given family have probably diverged too much from each other to cross-hybridize at the relatively moderate stringencies used here. The exception to this rule, C5, lends support to the above interpretation. It is the only V region that is conserved enough between species to hybridize well to squirrel, rabbit and human genomic DNAs, consistent with the observation that it is also the only V region to hybridize to multiple bands in the species surveyed³⁻⁸. The simplest interpretation of these hybridization patterns is that variation in some V_β sequences is under more stringent constraints than others, and that such V_β regions will therefore have retained enough sequence homology between duplicated copies in a given genome to appear as gene families. It also suggests that very low-stringency hybridization experiments may reveal that some of the other V_β sequences are also members of homologous gene families.

Homology with immunoglobulin

Based on amino acid homology, the published sequence of 86T1 more closely resembles a variable region from a κ chain (McPC 603, 33 of 84 amino acids homologous) than a variable domain from either a λ (MOPC 104e, 25 of 86) or heavy chain (93G7, 26 of 90)². Each of the additional V_β s reported here is more

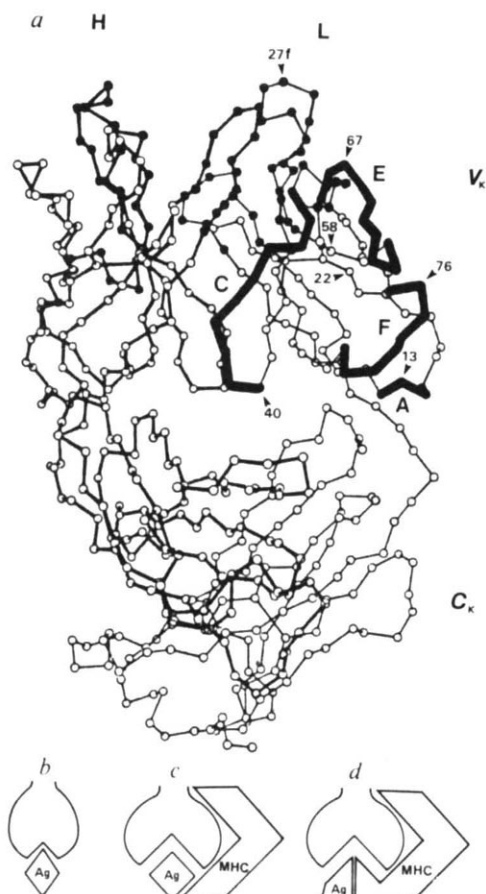


Fig. 6 Possible T-cell receptor structure and implications for recognition. *a*, An α -carbon skeleton of the X-ray crystallographic structure of McPC 603 (ref. 43). Thick lines denote heavy-chain residues. Closed circles correspond to classical hypervariable residues. The regions of sequence identified in Fig. 3*B* (A, C, E, F) that correspond to V_k framework regions, but are significantly more variable in V_β s, are shown as solid bars. *b*, A schematic diagram of antibody-antigen (Ag) interaction; *c*, *d*, models of T-cell receptor binding.

homologous with this κ V region than with these λ or heavy chains. Among the conserved residues that form the interface between V_H and V_L ⁴³, the T-cell receptor conserves the same amino acids at more of the light-chain contact residues (Tyr 36, Gln 38, Leu 46) than those of heavy chain (Gln 39, Leu 45). By these criteria, V_β most resembles a κ variable region. The only structural features which are more heavy chain-like are the number of amino acids from the invariant tryptophan to the last cysteine of the V region (4 or 5 residues shorter in V_κ than in V_β) and the fact that the last V_β framework region extends several amino acids past this cysteine, as in V_H . These similarities of V_β to V_H at the carboxy-terminal end may relate to the presence of D region elements^{4,5} and to the larger J regions characteristic of heavy chains^{4,6}. The intron-exon structure of the constant region with its separate hinge and transmembrane regions is also more heavy chain-like⁶.

Despite the similarities in amino acid sequence between V_β and V_κ , the relative lack of primary structure conservation among V_β s makes it important to determine whether the predicted secondary structure of these V_β s differs significantly from the known structure of V_κ and whether secondary structure is predicted to be conserved among V_β s. We therefore used the empirically derived algorithm of Chou and Fasman⁴² to locate regions of sequence predicted to be probable β -sheets or β -turns. We assumed that V_β s, like V_H and V_κ , have little or no α -helical content. Figure 5 shows the β -sheets and turns predicted by the algorithm of Chou and Fasman for six of the seven V_β s reported here. The predicted secondary structures of representative heavy- and light-chain V regions are included for comparison.

This analysis reveals that, among these V_β s, seven β -sheets are consistently predicted (five of six or better) in regions of sequence that correspond to β -sheets in immunoglobulin. The only apparent departure of the V_β predicted secondary structure from the known immunoglobulin structure is in residues 65–69 whose V_κ counterparts are known to comprise a β -sheet, but are predicted to be a sheet in only one of six V_β s (E1). The lack of a β -sheet in this region of sequence has precedent in immunoglobulin, as the Chou-Fasman algorithm often fails to predict the corresponding sheet in κ (unpublished observation). Thus, despite the extreme variability in primary structure between V_β s, the secondary structure is predicted to be conserved and to be immunoglobulin-like.

Models of T-cell receptor binding

Because of the similarities in the primary and predicted secondary structure between the T-cell receptor and the extensively characterized immunoglobulins, it is instructive to ask where, by analogy with immunoglobulin, putative new V_β hypervariable regions lie on the molecule. Historically, the assumption that regions of sequence which are highly variable among many immunoglobulins correspond to portions of the molecule which make physical contact with antigen has been a useful principle in analysing the antigen-binding site of immunoglobulins¹³. We have therefore applied this heuristic to the T-cell receptor β -chain in order to derive the implications for antigen-MHC interactions from the predicted location of putative new HV regions. Adapted from ref. 43 is an α -carbon skeleton of an immunoglobulin Fab from the myeloma McPC 603. Classical hypervariable regions are indicated on this Fab model by closed circles; regions of V_κ that correspond to putative new V_β hypervariable regions are indicated by heavy bars. Regions A and F correspond to external loops between antiparallel β -sheets, as do classical immunoglobulin HV regions. Region E corresponds to a β -turn and sheet that interacts with solvent in the V_κ molecule. Curiously, these regions lie external to the classical immunoglobulin antigen-binding site. Their variability in V_β suggests that important interactions may be occurring on the outer surface of the receptor, perhaps with polymorphic histocompatibility determinants. Region C corresponds to a β -turn and sheet that is highly conserved at the amino acid level in V_κ and makes multiple hydrophobic interactions and hydrogen bonds with the immunoglobulin heavy chain⁴³. In the T-cell receptor, despite the greater variation in amino acid residues in this region, the secondary structure is predicted to be highly conserved (Fig. 5). Although based on a small number of sequences, the variability of these residues in V_β suggests that they interact with solvent, antigen or the MHC; alternatively, the other chain of the T-cell receptor (α) has complementary variability in the residues in contact with this region of V_β .

Present cellular evidence indicates that a single T-cell receptor is responsible for recognizing both MHC and antigenic determinants^{44–46}. Assuming that the binding site(s) of this receptor consists of the α - and β -chains and that they fold in an approximately immunoglobulin-like fashion, the localization of the putative new hypervariable regions (Fig. 6*a*) suggests that the T-cell receptor may have the properties shown schematically in Fig. 6*c*, *d*. Specifically, it may have a larger binding pocket than that of immunoglobulin and its outer surface may be in contact with polymorphic MHC determinants. An alternative model (not shown) proposes the interaction of MHC with the classical immunoglobulin-like binding site, and antigen interacting with the putative new HV regions on the outer surface of the receptor. These models predict that different MHC restriction elements could alter the antigen-binding properties of the β -chain through allosteric effects. Preliminary evidence with anti-cytochrome c-reactive hybridomas is consistent with these models, as different β -chain V regions are used against the same antigenic peptide in the context of different MHC restricting alleles (S. Hedrick, unpublished observation).

We thank Dr John Bell for his generous gift of human DNA, Drs Robert Schimke and Chuck Gasser for CHO DNA, Dr C.

Benoist for thymocyte cDNA libraries, Dr Tim Mosmann for LB2 cells and Dave Austin for the use of his immunoglobulin model. We also thank Dr Gordon Freeman for preparing E1 and C5 mRNA and for assistance in the cDNA cloning, Dr Steve Crews for S107 DNA and critical reading of the manuscript, Dr Elvin Kabat for critical remarks on the manuscript,

Dr Harvey Cantor for helpful discussions, and Kathy Redman and Gloria Samuel for help in preparing the manuscript. P.A.P. is supported by a NIH predoctoral training grant in molecular and cellular biology. Much of the work described here was funded by grants to M.M.D. from the NIH and the American Cancer Society.

Received 31 August; accepted 20 September 1984.

- Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149–153 (1984).
- Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153–158 (1984).
- Yanagi, Y. *et al.* *Nature* **308**, 145–149 (1984).
- Chien, Y., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322–326 (1984).
- Siu, G. *et al.* *Cell* **37**, 393–401 (1984).
- Gascoigne, N. R. J., Chien, Y., Becker, D. M., Kavalier, J. & Davis, M. M. *Nature* **310**, 387–391 (1984).
- Kavalier, J., Davis, M. M. & Chien, Y. *Nature* **310**, 421–423 (1984).
- Malissen, M. *et al.* *Cell* **37**, 1101–1110 (1984).
- Siu, G. *et al.* *Nature* **311**, 344–350 (1984).
- Acuto, O. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 3851–3855 (1984).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51–177 (1979).
- Bevan, M. & Fink, P. *Immun. Rev.* **42**, 3–19 (1978).
- Wu, T. T. & Kabat, E. A. *J. exp. Med.* **132**, 211–250 (1970).
- Fresno, M., Nabel, G., McVay-Boudreau, L., Furthmayer, H. & Cantor, H. *J. exp. Med.* **153**, 1246–1259 (1981).
- Clayberger, C., Dekruyff, R. H., Aisenberg, J. & Cantor, H. *J. exp. Med.* **157**, 1906–1919 (1983).
- Giedlin, M., Longenecker, B. & Mosmann, T. (submitted).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–170 (1978).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **3**, 280–289 (1983).
- Davis, M. M., Chien, Y., Gascoigne, N. R. J. & Hedrick, S. M. *Immun. Rev.* **81**, 233–255 (1984).
- Rao, A., *et al.* *Cell* **36**, 879–888 (1984).
- Hedrick, S. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).

- Maxam, A. & Gilbert, W. *Methods Enzym.* **65**, 499–560 (1980).
- Saito, H. *et al.* *Nature* **309**, 757–762 (1984).
- Chien, Y., Kavalier, J. & Davis, M. M. *Proc. natn. Acad. Sci. U.S.A.* (submitted).
- Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *Sequences of Immunological Interest* (US Dept of Health and Human Services, Washington, D.C., 1983).
- Nakanishi, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 6984–6988 (1982).
- Kraig, E. *et al.* *J. exp. Med.* **158**, 192–209 (1983).
- Kronenberg, M. *et al.* *J. exp. Med.* **158**, 210–227 (1983).
- Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D. M.) (IRL, Oxford, 1984).
- Melchers, I. *J. exp. Med.* **156**, 1587–1603 (1982).
- Hamann, U. *et al.* *J. Immun.* **130**, 7–14 (1983).
- Capra, J. D. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **71**, 845–848 (1974).
- Litman, G. *et al.* *Nature* **303**, 349–352 (1983).
- Gojobori, T. & Nei, M. *Molec. Biol. Evol.* **1**, 195–212 (1984).
- Crews, S. *et al.* *Cell* **25**, 59–66 (1981).
- Dickerson, R. & Geis, I. *Hemoglobin: Structure, Function, Evolution, and Pathology* (Benjamin/Cummings, Menlo Park, 1983).
- Romer, A. *Vertebrate Paleontology* (University of Chicago Press, Chicago, 1966).
- Litman, G. *et al.* *Nucleic Acids Res.* **10**, 3371–3380 (1982).
- Bentley, D. *et al.* *Cell* **24**, 613–623 (1981).
- Brodeur, P. & Riblet, R. (in preparation).
- Cory, S. *et al.* *J. molec. appl. Genet.* **1**, 103–116 (1981).
- Chou, P. & Fasman, G. A. *Rev. Biochem.* **47**, 251–276 (1978).
- Davies, D. & Metzger, H. A. *Rev. Immun.* **1**, 87–117 (1983).
- Kappler, J. W., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198–1214 (1981).
- Heber-Katz, E. *et al.* *J. exp. Med.* **155**, 1086–1099 (1982).
- Hunig, T. R. & Bevan, M. J. *J. exp. Med.* **155**, 11–125 (1982).
- Maniatis, T., Fritsch, E. & Sambrook, J. in *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

LETTERS TO NATURE

Detection of possible optical flashes from the γ -ray burst source GBS0526–66

H. Pedersen*, J. Danziger*, K. Hurley†, G. Pizzichini‡, C. Motch§, S. Ilvasky§, N. Gradmann||, W. Brinkmann||, G. Kanbach||, E. Rieger||, C. Reppin||, W. Trumper|| & N. Lund¶

* European Southern Observatory, Karl Schwarzschild-Strasse 2, D 8046 Garching bei München, FRG

† Centre d'Etude Spatiale des Rayonnements, BP 4346, 31029 Toulouse Cedex, France

‡ Istituto TESRE/CNR, Via dei Castagnoli, 1, 40126 Bologna, Italy

§ Observatoire, 41 bis, Avenue de l'Observatoire, 25000 Besançon, France

|| Max Planck Institut für Physik und Astrophysik, Institut für Extraterrestrische Physik, D 8046 Garching bei München, FRG

¶ Danish Space Research Institute, Lundtoftevej 7, DK 2200 Lyngby, Denmark

The unidentified source of the γ -ray burst of 5 March 1979, GBS0526–66, may be located in the supernova remnant N49 in the Large Magellanic Cloud¹, though some authors^{2,3} have questioned the implied distance of 55 kpc. Among several unusual features of the source (reviewed by Cline⁴) is its repeatability: a total of 16 γ -ray bursts have been reported³, which may be periodic⁵. Only γ -ray bursts have been detected from this source. The situation for other γ -ray burst sources is not very different. Optical transients are known, in three cases^{6,7}, to have occurred near positions of later γ -ray bursts. Even so, no source has been identified optically^{8–10}, and only upper limits are known for simultaneous emission of optical radiation with γ -ray bursts¹¹. Here we report that GBS0526–66 was monitored for ~910 h, using a high speed photoelectric photometer, attached to a 50 cm telescope at ESO/La Silla. Three short optical flashes were found, which may possibly be related to the γ -ray burst phenomenon. The most intense of these occurred on 1984 February 8 UT 07:42. It reached a maximum brightness corresponding to $m(\text{visual}) = 8.7$. The shape of the light curve bears remarkable similarity to the initial 200 ms of the 1979 March γ -ray burst.

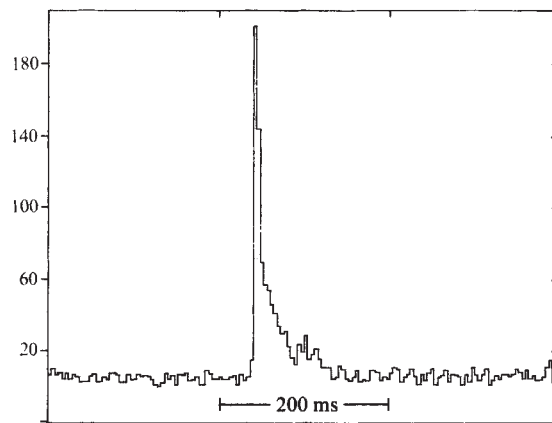


Fig. 1 An event observed 1983 October 27 UT 01:28. The time bin is 4 ms. Vertical axis corresponds to no. of counts, uncorrected for dark count or sky.

Photometric observations of the 1979 March 5 γ -ray burst error box were made during the period 1983 October 5 to 1984 February 23. A single channel photometer was used with a rectangular entrance diaphragm, 20 by 80 arc s, aligned with the error box. Offset-guiding allowed the diaphragm to remain centred on the error box during almost all the observations and, in particular, during the burst moments discussed here. Most measurements were done through a Johnson V filter ($\lambda = 550$ nm, $\Delta\lambda = 80$ nm). During the period up to 1983 October 20, an EMI 6256 photomultiplier was used; later data were acquired with an RCA C31034A tube. Pulse counts were written on magnetic tape at equal time intervals. The time bin was 4 ms until 1983 November 26, and 1 ms afterwards. Integrations with a time-base of 4 ms were kept in floating-point words; those of 1 ms were packed with two integers per 16-bit word. The resulting maximum count rate $255,000 \text{ s}^{-1}$ corresponds to $m(\text{vis}) = 6.6$. A strip-chart recorder allowed immediate inspection of the count rate. The observer could set the time resolution, usually 1 or 0.5 s. During part of the observations, an on-line burst detecting procedure informed the observer about possible bursts in order that the sky conditions could be checked. An atomic beam standard frequency clock was used for time keeping. Absolute

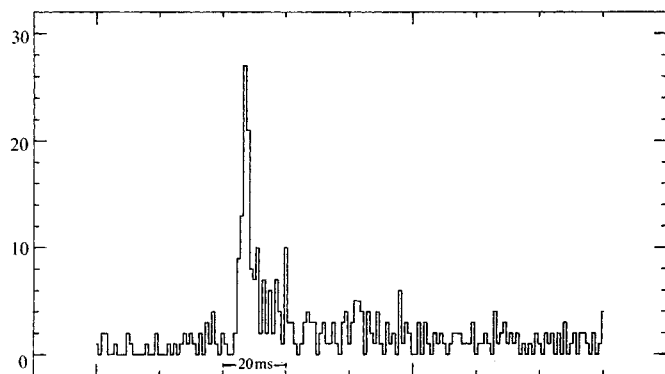


Fig. 2 An event observed on 1984 January 30 UT 03:20. The time bin is 1 ms.

time is presently known only to ± 100 ms because of an error in the calibration procedure. Data from 1983 November 2 to 19 were not recorded digitally, but only using the strip-chart recorder. The constancy of the system sensitivity was checked against the nearby star SAD249271, whereas photometric calibration was done using data from various E-region stars taken from ref. 12. The observations were carried out throughout the moon cycle, in spite of occasional light cloud cover.

A number of natural and man-made phenomena can cause spurious burst detection. Some, including electronic switching noise, lightning, the xenon flashlamps on aircraft, and meteors passing through the field of view, are of only millisecond duration, or less. We believe we have eliminated such effects by disregarding events lasting < 20 ms. Longer events, which may possibly mimic cosmic bursts, include the decay of luminescent meteor trains in the field of view, diffuse sky background illumination due to flaring fireballs, and the passage of artificial Earth satellites across the field of view. Little appears to be known about the first of these effects. Data published by Sekanina¹³, indicate that fireball light curves often have multiple maxima, and decrease in luminosity more rapidly than they increase, which is inconsistent with the events observed here. Slow-moving, artificial satellites would generally produce symmetrical light curves. Occasionally, a satellite may change brightness while moving only a few arcseconds, with unpredictable photometric results. Symmetric events of 32–90 ms duration have been observed on ten occasions, but we have not been able to demonstrate coincidence with a satellite passage. Finally we mention that Moon-illuminated clouds and contrails from jet aircraft can produce 'bursts' of a wide range of shapes and different durations. Contrails are particularly bothersome, as they may occur under all photometric conditions.

All data acquired during the 910 h observation period have been examined for the occurrence of highly significant events. The following three events comprise all flashes not suspected of being man-made, and of > 20 ms duration. Events discarded include the ten possible satellite passages, and a few chaotic events possibly attributable to low line-voltage. All three events were observed under perfect conditions—dark and clear sky; diaphragm well centred on error box. Their light curves are shown in Figs 1–3.

(1) 1983 October 27 UT 01:28 (Fig. 1): The burst was registered by the triggering routine. The two observers in the dome noted normal sky conditions. The rise time, half-width (FWHM) and total duration were < 4 , 8, and 90 ms, respectively. Maximum brightness was $m(\text{vis}) = 8.4$. The amount of energy in the first 4 ms time bin, relative to that in the tail, was $\sim 1/3$. This is possibly inconsistent with the detection of a meteorite passage, followed by the decay of its train. However, this explanation cannot be precluded. It is also possible that the rise time is in fact resolved: the bin preceding the maximum may indicate the start of the phenomenon. Alternatively it could be a statistical fluctuation.

(2) 1984 January 30 UT 03:20 (Fig. 2): The telescope was unattended. The rise time, half-width and total duration were

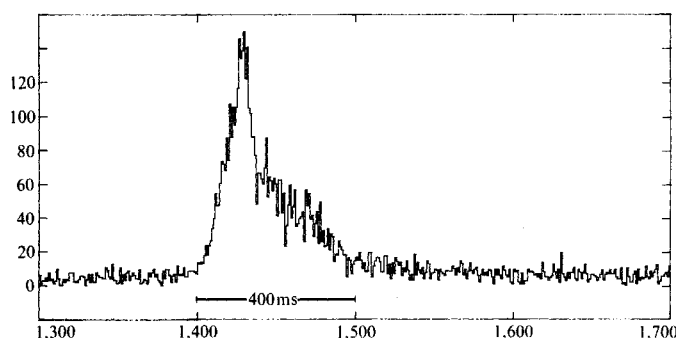


Fig. 3 An event observed on 1984 February 8 UT 07:42, at phase zero of the possible 164-day period determined by Rothschild and Lingenfelter⁵. The time bin is 4 ms.

3, 3, and 60 ms, respectively. The maximum brightness was $m(\text{vis}) = 9.9$.

(3) 1984 February 8 UT 07:42 (Fig. 3): The telescope was unattended. Maximum brightness was $m(\text{vis}) = 8.7$. The rise time, half-width and total duration were 110, 90 and 1,700 ms respectively. The event differs from the others by its much longer duration. Its resolved rising edge is a strong argument against a meteor explanation. The maximum brightness, integrated over the visible band, and referred to the top of the atmosphere, is $1.0 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2}$. The corresponding fluence is $1.6 \times 10^{-10} \text{ erg cm}^{-2}$.

The simultaneous detection of γ -ray bursts and optical flashes would convincingly establish their cosmic origin. During our observations, four spacecraft with γ -ray detectors were in operation: ICE, PVO, SMM and Prognoz 9. Preliminary analysis of these has revealed no response at the time of any of the optical flashes. The detector thresholds depend strongly on the time history and energy spectrum of a burst, but can be conservatively characterized by a fluence of $10^{-6} \text{ erg cm}^{-2}$, at energies above 100 keV. A pending analysis of real-time data may reduce this value by a factor of two. The absence of γ -ray burst detections is not surprising, given the thresholds and sensitivities of the satellites: none detected any of the soft, repeating bursts observed by the KONUS instruments. The inferred fluence ratio $E(\text{opt})/E(\gamma) > 1.6 \times 10^{-4}$ is consistent with similar, published data: Grindlay *et al.*¹¹ found the upper limit 10^{-3} for one particular γ -ray burst, which could not be detected optically, and Schaefer^{6,7} deduced values near 10^{-3} for the three historic, optical transients mentioned above.

In spite of the absence of γ -ray detections, several facts make the optical events remarkable: (i) Two of them have fast rise times and relatively slow decay, reminiscent of the 1979 March 5 event and of some of the recurrent bursts. (ii) The shape of the third burst is similar to the initial, impulsive phase of the 1979 March 5 γ -ray burst¹⁴. Both phenomena show a plateau at about half intensity, following the peak. The main difference is the finite rise time of the optical flash and its longer duration. (iii) The moment the third event occurred is consistent with the day of phase zero, as extrapolated from the formula given by Rothschild and Lingenfelter⁵: at this phase, a rather high burst rate may be expected.

We note that all published light curves of γ -ray bursts from GBS0526–66 (refs 1, 15–17) show very short rise times. If the 1984 February 8 optical flash represents reprocessed γ -rays from such rapid events, the rise time and half-width of the optical flash may be used to estimate a typical dimension of 30,000 km for the emitter. Unless the γ -ray bursts are strongly beamed, this seems to rule out reprocessing in the atmosphere of a (non-degenerate) neighbour star as proposed in the model of London and Cominsky¹⁸. However, it is consistent with small-disk models and with models involving cyclotron reprocessing. The latter have been considered by Woosley¹⁹.

We thank R. Lindskov-Simonsen and R. Castillo for help in making observations and data analysis; S. Laursen for work on hardware and software; the Danish Board for Astronomical

Research for a travel grant; and G. Andreoni for work with telescope tracking models.

Received 23 August; accepted 5 October 1984

1. Cline, T. L. *et al. Astrophys. J. Lett.* **255**, L45–L48 (1982).
2. Zdziarski, A. A. in *Accreting Neutron Stars* (eds Brinkmann, W. & Trumper, J.) 246–249 (Garching, 1982).
3. Golenetskii, S. V., Ilyinskii, V. N. & Mazets, E. P. *Nature* **307**, 41–43 (1984).
4. Cline, T. L. *Comments Astrophys.* **9**, 13–21 (1980).
5. Rothschild, R. E. & Lingenfelter, R. E. *Nature* (submitted).
6. Schaefer, B. E. *Nature* **294**, 722–724 (1981).
7. Schaefer, B. E. *et al. National Aeronautics and Space Administration Preprint No.* 86126 (1984).
8. Chevalier, C. *et al. Astr. Astrophys. Lett.* **100**, L1–L3 (1981).
9. Pedersen, H. *et al. Astrophys. J. Lett.* **270**, L43–L47 (1983).
10. Schaefer, B. E. *et al. Astrophys. J. Lett.* **270**, L49–L52 (1983).
11. Grindlay, J. E., Wright, E. L. & McCrosky, R. E. *Astrophys. J. Lett.* **192**, L113–L114 (1974).
12. Graham, J. A. *Publ. astr. Soc. Pacif.* **94**, 244–265 (1982).
13. Sekanina, Z. *Astr. J.* **88**, 1382–1414 (1983).
14. Baan, W. A. *Astrophys. J. Lett.* **261**, L71–L75 (1982).
15. Golenetskii, S. V. *et al. Sov. Astr. Lett.* **5**, 340–342 (1979).
16. Mazets, E. P., Golenetskii, S. V., Il'inskii, V. N., Aptekar, R. L. & Guryan, Yu. A. *Nature* **282**, 587–589 (1979).
17. Cline, T. L. *et al. Astrophys. J. Lett.* **237**, L1–L5 (1980).
18. London, R. A. & Cominsky, L. R. *Astrophys. J. Lett.* **275**, L59–L63 (1983).
19. Woosley, S. E., *Conf. Proc. Am. Inst. Phys.*, **115**, 485–511 (1984).

Is Geminga a very close neutron star binary?

P. E. J. Nulsen & A. C. Fabian

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Geminga (2CG195+4) is one of the brightest sources in the sky at photon energies around 1 GeV (refs 1–3). It has been tentatively identified⁴ with an unusual soft X-ray and optical object for which the γ -ray-to-X-ray-to-optical luminosity ratios are $\sim 10^6:10^3:1$, which may now have been confirmed by the discovery⁵ of a ~ 60 -s periodic variation in the X-ray source, similar to that reported in γ rays. The period has been increasing over the past 10 yr at a rate such that it doubles in ~ 400 – 800 yr. The implied rotational energy loss would supply the γ -ray luminosity of Geminga if the spinning object were a neutron star at a distance of ~ 10 pc, consistent with the very low observed X-ray absorption. A lone pulsar with a spin period of ~ 60 s cannot spin down at the required rate. The observed period would need to be due to precession or some other process. We show here that the presence of a close neutron star companion provides a suitable mechanism. All neutron star binaries may evolve through a Geminga-like phase.

A non-corotating conductor orbiting in a pulsar magnetosphere creates a circuit in which currents may flow along the field lines and through the neutron star (compare with the unipolar inductor of ref. 6). The currents will be driven by, and will tap energy from, the rotation of the pulsar. The optical and X-ray luminosities of Geminga are so small that plausible candidates for a conductor are severely restricted. A normal stellar companion would have to orbit at too great a distance and an ionized accretion disk (compare with ref. 7) or a white dwarf would be too luminous. We consider here that the conductor is another neutron star with a weaker, opposed, magnetic field. The 60-s period is identified with the difference between the spin frequency, F , of the neutron star (primary; with the larger magnetic moment, $\mathcal{M}_1$) and the orbital frequency of the other neutron star (secondary of magnetic moment $\mathcal{M}_2$). This orbital period is several minutes.

The common flux between two opposed coplanar dipoles of separation a is

$$\Phi = \frac{\mathcal{M}_1}{a} \alpha (1 - \alpha^2)^2 \int_0^{2\pi} \frac{\cos \phi \, d\phi}{(1 + \alpha^2 - 2\alpha \cos \phi)^{3/2}} \\ = 3\pi\alpha^2 \frac{\mathcal{M}_1}{a} f(\alpha) \quad (1)$$

where $\mathcal{M}_2 = \alpha^3 \mathcal{M}_1$. $f(\alpha)$ varies smoothly from 0.85 to 1 as α varies from 1 to 0. Hereafter we adopt $f(\alpha) = 1$. The effective cross-section of the secondary ($3\pi\alpha^2 a^2$ for small α) may greatly

exceed its geometric cross-section. The structure of the magnetic field is outlined in Fig. 1. The maximum voltage around the generator circuit occurs between the pair of 'last common field lines' coplanar with the two dipoles and the rotation axis. By mapping the field lines onto the primary, we obtain for small α

$$V_{\max} = 2\sqrt{3} \alpha \mathcal{M}_1 (F - \sqrt{GM/a^3})/a \\ \approx 3.5 \times 10^{12} \alpha \mathcal{M}_{21} a_8^{-1} \text{ V} \quad (2)$$

where $M = (M_1 + M_2)$ is the mass of the system, $a = 10^8 a_8$ m and $\mathcal{M}_1 = 10^{21} \mathcal{M}_{21} \text{ T m}^3$. A magnetic moment $> 2.5 \times 10^{21} \text{ T m}^3$ takes the surface field beyond the quantum limit where spontaneous pair creation is expected. A further small voltage difference of $\frac{3}{2} \alpha^2 \mathcal{M}_1 \omega'/a$ can occur if the spin of the secondary deviates from corotation by ω' .

The current flow in the magnetosphere cannot produce magnetic fields in excess of the dipolar ones^{8,9}, which in the present case limits the total current to less than

$$I_{\max} \approx 2\pi\alpha \mathcal{M}_1 / \mu_0 a^2 \\ = 5.0 \times 10^{11} \alpha \mathcal{M}_{21} a_8^{-2} \text{ A} \quad (3)$$

Deep within the crust of a neutron star, the electrical conductivity is isotropic and extremely large, but it is anisotropic at the surface if the magnetic moment is large^{10,11}. This region dominates the effective stellar resistance. A crude upper limit is obtained by applying a value for the surface conductivity ($\sigma_1 \approx 10^9 \text{ S m}^{-1}$, ref. 11) to the whole star;

$$R_{\text{eff}} < \frac{4a}{3\pi\alpha^2 r_*^2 \sigma_1} \\ \approx 4 \times 10^{-10} \alpha^{-2} a_8 \Omega \quad (4)$$

where r_* is 10 km. Since

$$R_{\text{eff}} \ll R_{\min} = V_{\max} / I_{\max} \\ = \frac{\sqrt{3} \omega a}{\pi c} Z_0 \quad (5)$$

where $\omega = F - \sqrt{GM/a^3}$ and $Z_0 = \sqrt{\mu_0/\epsilon_0}$ ($\approx 38 \Omega$) the current is self-limited in the magnetosphere.

The power dissipated in the magnetosphere

$$P_m \approx V_{\max} I_{\max} = 4\sqrt{3} \pi \alpha^2 \mathcal{M}_1^2 \omega a^{-3} \mu_0^{-1} \\ = 1.7 \times 10^{24} \alpha^2 \mathcal{M}_{21}^2 \omega_{-1} a_8^{-3} \text{ W} \quad (6)$$

where $\omega = 0.1 \omega_{-1}$. It can be heuristically derived as the rate of working of the magnetic stress in the common flux tube. P_m should equal the power associated with the spin-down of the primary neutron star

$$P_R \approx 3.2 \times 10^{24} \text{ W} \quad (7)$$

for a period derivative of $3 \times 10^{-9} \text{ s s}^{-1}$, mass of $1 M_\odot$ and moment of inertia of $6 \times 10^{37} \text{ kg m}^2$. We propose that the electric potential ($\sim V_{\max}$) accelerates particles that radiate the observed γ rays. Synchrotron and curvature radiation from 1 to 10 TeV electrons can produce γ radiation in the appropriate range. Whatever the emission mechanism, its efficiency for γ -ray production must be very high ($> 99\%$), comparable with that of the pulsed emission from the Vela pulsar¹². We note that the large opacity to pair production on the magnetic field prevents γ rays emitted near the neutron star poles from escaping¹³. Ignoring beaming, the observed flux¹⁻³ of $2.4 \times 10^{-12} \text{ W m}^{-2}$ requires Geminga to be at a distance of ~ 11 pc.

A completely charge-separated beam will limit the current to $< I_{\max}$, since the potential variation across a current beam made of electrons or ions alone is roughly $\delta V \approx I/\epsilon_0 v_D > IZ_0$, where the drift velocity $v_D < c$. To attain P_m , we require the maximum current to flow and thus nearly neutral beams. This is achieved if the incoming particles dislodge charges from the neutron star surfaces. Pair production by γ rays in the intense magnetic field near to the neutron star may also supply a substantial fraction of the current carriers¹³.

The flux tube is an ellipse at the primary of semi-axes of $\sqrt{3/2}\alpha\sqrt{r_*^3/a}$ and $\sqrt{3}\alpha\sqrt{r_*^3/a}$, and for the parameters that give P_R its area, $A_\Phi \approx 4.7 \times 10^4 \alpha^m a_8^{-1} \text{ m}^2$, where $m = +2$ and -1 on the primary and secondary respectively. If the X-ray flux is due to black-body emission, we derive an area of $\approx 5 \times 10^4 T_6^{-2} \text{ m}^2$, where $T = 10^6 T_6 \text{ K}$. The X-ray observations⁴ suggest that the black-body temperature $T_6 \approx 1$. The optical flux can also originate here; the precise optical-to-X-ray luminosity ratio depending on the observing windows but could be $\ll 10^{-3}$. Further emission may originate at the secondary and in the beams themselves. The work function of the surface is $\sim 1\text{--}10 \text{ keV}$ at the relevant magnetic field strengths¹⁰, so thermionic emission at the black-body temperature cannot supply the neutralizing charge, although there may be an optically-thin outer region heated to greater temperatures.

The γ -ray modulation is assumed to be due to variation in the magnetic field strength at the base of the common flux tube as it travels over the primary star, caused by misalignment of the rotation and magnetic axes and/or non-dipolar components of the field. The 60-s period is identified as $2\pi/\omega$ (see equation (5)). There are three independent frequencies in this system (two spins and the orbit) which could lead to great complexity in its periodic behaviour. In particular, the frequency of variation may be different in different wavebands, so that the period derivative from Bignami *et al.*⁵ could even be incorrect.

In order that $P_m \approx P_R$, the orbital separation of the two neutron stars, a , must be $\leq 10^8 \text{ m}$. If they are each of $1 M_\odot$, then the orbit will decay due to gravitational radiation in $\approx 30,000 \text{ yr}$. The power radiated by gravitational waves is $\approx 3 \times 10^{29} a_8^{-5} \text{ W}$ and exceeds that as electromagnetic radiation. In this model, Geminga is essentially a 'gravity wave binary'. The relative strain at Earth is $\approx 6 \times 10^{-20} a_8^{-1}$ and is unfortunately as yet undetectable¹⁴.

In view of the short anticipated life of Geminga, both to spin-down ($\sim 800 \text{ yr}$) and to gravitational radiation ($\approx 30,000 \text{ yr}$), it is important to consider its formation. Binary neutron stars may form in the Galaxy at a rate¹⁵ of $\sim 3 \times 10^{-4} \text{ yr}^{-1}$ with orbital periods of between one and a few hours. The *a priori* probability of the existence of such a neutron star binary within $\sim 100 \text{ pc}$ approaches one, as the lifetime to gravitational radiation is $\sim 10^7\text{--}10^8 \text{ yr}$. To observe it during a phase lasting $\leq 10^3 \text{ yr}$ and at a distance of $\sim 10 \text{ pc}$, however, makes Geminga exceptional.

Beaming of the γ -ray flux can relax the distance estimate. We also note that any model involving a rapidly-spinning neutron star in which the 60-s period is due to free precession and which spins down in $< 10^3 \text{ yr}$ dissipates $> 10^{30} \text{ W}$. It is then difficult to see how Geminga could be at a distance of $< 100 \text{ pc}$. A possible alternative interpretation in which the object is nearby would be that the $\sim 800 \text{ yr}$ time scale is due to change in distortion of such a pulsar with a weak magnetic field and spin period of $\sim 100 \text{ ms}$. The free-precession frequency is proportional to the difference between the principal moments of inertia. It is then a mystery as to why Geminga is not a radio pulsar⁴. The object would also be a strong source of gravitational waves.

After formation, a neutron star binary goes through a pulsar phase until the speed-of-light cylinder of the star with the greater magnetic moment passes beyond the orbital separation that is

$$F \leq c/a \quad (8)$$

After this phase the unipolar generator mechanism will take over control of the spin-down rate, so that from equation (6)

$$\frac{dF}{dt} \approx -\frac{4\sqrt{3}\pi\alpha^2 M_1^2 (F - \sqrt{GM/a^3})}{\mu_0 \mathcal{I} F a^3} \quad (9)$$

where $\mathcal{I}$ is the moment of inertia of the primary. The spin-down has an insignificant effect on the orbit (because the orbital angular momentum always greatly exceeds that in the spin), so that the orbit decays due to gravitational radiation¹⁶ as

$$a = \left[a_0^4 - \frac{256 G^3 M_1 M_2 M}{5 c^5} t \right]^{1/4} \quad (10)$$

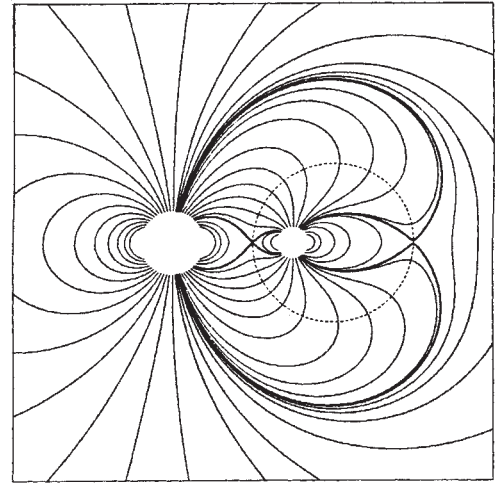


Fig. 1 Magnetic field of a pair of opposed coplanar dipoles. This shows some field lines (not equally spaced in flux) in the plane containing the two dipoles, for $\alpha = 0.5$ (see equation (1)). The boundary of the common flux tube, the 'last common field lines', is shown as bolder than the rest. The dashed circle separates the regions dominated by the two dipoles. The magnetic field is zero on a circle surrounding the weaker dipole where the last common field lines cross over. Although distorted, this feature remains when the dipoles are (moderately) misaligned and when currents $\approx I_{\text{max}}$ flow. There are large potential differences across the boundary of the common flux tube, particularly near the cross-over points where the two regions of isolated field come into contact.

from the initial separation a_0 . While the spin frequency, $F \gg$ the orbital frequency, $\sqrt{GM/a^3}$, equations (9) and (10) may be combined to give

$$(1 - F/F_0) \approx \Gamma(1 - a/a_0) \quad (11)$$

where F_0 is the initial spin frequency, and

$$\Gamma \approx 210 \alpha^2 \chi^2 / \xi a_0^2 \quad (12)$$

where $a_0 = 10^9 a_9 \text{ m}$, $F = \xi c/a_0$ (compare with equation (8)), the polar magnetic field on the primary is

$$B_1 = \chi m_e^2 c^2 / e \hbar = 4.7 \times 10^9 \chi \text{ T} \quad (13)$$

and the primary and secondary stars have both been given the parameters adopted above. The remarkable feature of equation (12) is that for reasonable values of the parameters $\Gamma \sim 1$. The binary orbit thus decays at a comparable rate to the spin of the primary. Failing this condition, either the orbit would decay without any significant spin-down ($\Gamma \ll 1$; mostly spin-up) or the spin-down would bring the primary close to corotation before the orbit decayed significantly ($\Gamma \gg 1$). (It is possible that the system did start with $\Gamma \ll 1$, but in that case the gravitational decay time now must be very short indeed, $\ll 800 \text{ yr}$, and the primary star must have started with a spin period $\approx 60 \text{ s}$.) The required high magnetic moment of the primary neutron star implies that it cannot be very old ($\leq 10^7 \text{ yr}$) if current ideas¹⁷ on the decay of magnetic fields are correct.

There is an alternative to the above evolutionary picture which can also produce the Geminga system. The unipolar generator is self exciting once a substantial current has been established because this heats the foot of the flux tube as discussed above. However, if there is very little matter in the magnetosphere the currents may not be able to turn on spontaneously. Perhaps the event¹⁸ of AD 437 was associated with the start of the current flow, caused by accretion of matter passing through the system or by a crustquake¹⁹.

The final outcome of the close neutron star binary we have proposed for Geminga is uncertain²⁰, but it could be a black hole or a millisecond pulsar¹⁵.

We thank K. A. Arnaud, G. F. Bignami, D. O. Gough, E. Priest, J. E. Pringle and M. J. Rees for discussions.

Received 5 July; accepted 22 August 1984.

1. Fichtel, C. E. *et al. Astrophys. J.* **198**, 163–182 (1975).
2. Thompson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252–262 (1977).
3. Swanenburg, B. N. *et al. Astrophys. J. Lett.* **243**, L69–L73 (1981).
4. Bignami, G. F., Careveo, P. A. & Lamb, R. C. *Astrophys. J. Lett.* **272**, L9–L13 (1983).
5. Bignami, G. F., Careveo, P. A. & Paul, J. A. *Nature* **310**, 464–469 (1984).
6. Goldreich, P. & Lynden-Bell, D. *Astrophys. J.* **156**, 59–78 (1969).
7. Michel, F. C. & Dessler, A. J. *Astrophys. J.* **251**, 654–664 (1982).
8. Lamb, F. K., Aly, J.-J., Cook, M. C. & Lamb, D. Q. *Astrophys. J.* **274**, L71–L75 (1983).
9. Parker, E. N. *Cosmic Magnetic Field* (Clarendon, Oxford, 1979).
10. Flowers, E. G. *et al. Astrophys. J.* **215**, 291–301 (1977).
11. Itoh, N. *Mon. Not. R. astr. Soc.* **173**, 1P–3P (1975).
12. Bignami, G. F. & Hermesen, W. A. *Rev. Astr. Astrophys.* **21**, 67–108 (1983).
13. Arons, J. & Scharlemann, E. T. *Astrophys. J.* **231**, 854–879 (1979).
14. Anderson, J. D. *et al. Nature* **308**, 158–160 (1984).
15. Henrichs, H. F. & Van den Heuvel, E. P. J. *Nature* **303**, 213–216 (1983).
16. Clark, J. P. A. in *Sources of Gravitational Radiation* (ed. Smarr, L.) 447–459 (Cambridge University Press, 1979).
17. Blandford, R. D., Applegate, J. H. & Hernquist, L. *Astrophys. J.* **204**, 1025–1048 (1983).
18. Stephenson, F. R. Q. *Jl. R. astr. Soc.* **17**, 121–138 (1976).
19. Ellison, D. C. & Kazanas, D. *Astr. Astrophys.* **128**, 102–109 (1983).
20. Clark, J. P. A. & Eardley, D. M. *Astrophys. J.* **215**, 311–322 (1977).

Is Cygnus X-3 a monoenergetic 10^{17} eV accelerator?

A. M. Hillas

Physics Department, University of Leeds, Leeds LS2 9JT, UK

Like many other X-ray sources, Cygnus X-3 seems to be a binary system that includes a compact object, but it is unusual in being enshrouded in gas, in displaying occasional huge outbursts and in emitting much more energy above 10^{11} eV than in the keV X-ray region. If the compact object is a pulsar, the ultra-high-energy photons can be products of an accelerated beam of protons (or nuclei) striking the companion's atmosphere^{1,2}. It is noted here that a roughly monoenergetic spray of $\sim 10^{17}$ eV protons at $\sim 10^{39}$ erg s⁻¹, generating an electron-photon cascade in the environs of the companion star, best accounts for all photons above a few GeV, and might generate the gas shroud, as beam energy deposition inside the star could lift off a thick layer of matter to generate a fountain under the pulsar. This evidence on the nature of the accelerated beam is particularly significant as Cygnus X-3 seems likely to be responsible for most of the cosmic ray particles around 10^{17} eV which are at present being injected into the Galaxy.

γ Rays are seen briefly at two portions of the 4.8-h binary period, at phases reported variously as 0.15–0.25 and 0.63–0.8, taken from the time of deepest eclipse of the X-ray source; the spectrum extends (roughly as $E^{-2} dE$, where E represents energy) to 10^{16} eV (refs 3, 4). Vestrand and Eichler^{1,2} have suggested that we see the pulsar through the tenuous outer layers of the companion just before and after eclipse, and see photons from nuclear interactions (via π^0 decay) in the gas. This is supported by the finding that although a potential drop of 10^{17} V required to accelerate charged particles could hardly be found more than ~ 100 km from the pulsar, the 10^{16} eV photons cannot penetrate a transverse magnetic field exceeding 500 G (ref. 5) and hence must be formed far away, somewhere near the companion star ($\sim 5 \times 10^{10}$ cm distant), where the field is believed to be of this magnitude^{5,6}. However, three problems arise: (1) The diameter of the relevant outer surface of the companion must be very large¹, as it appears to eclipse the source of X rays and low-energy γ rays for $\sim 40\%$ of the orbit. Modelling of the intensity curves⁷ indicates that the shadow must be cast by gas present far outside any plausible Roche lobe of the companion unless the pulsar is extraordinarily light. (2) How is the spectrum generated? (3) How do the 10^{12} eV photons remain well collimated if they are derived from 10^{13} eV protons, as has generally been supposed? To explain radio emission, Vestrand⁶ requires a magnetic field (B) of 500–1,000 G to be embedded in the gas, and Stephens⁵ showed that such a field could cause the sharp end to the photon spectrum. The radius of gyration of 10^{13} eV protons would then be less than 10^{-2} of their path either before or after entering the gas, destroying all directionality. I will initially assume that this local field is present, but even if instead the only field arises from the pulsar, one still expects deflection

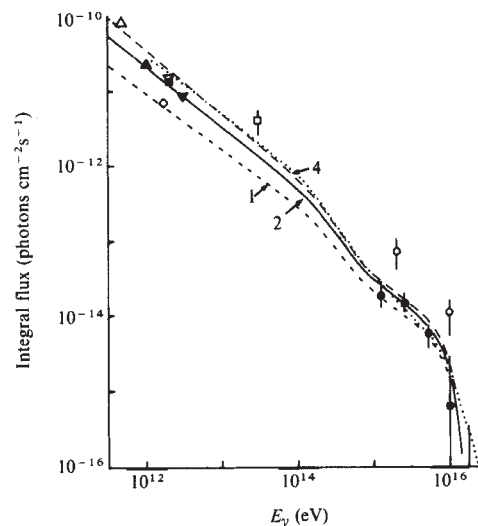


Fig. 1 Calculated photon spectrum resulting from cascades initiated by 10^{17} eV protons in gas surrounding the companion star (magnetic field present), for cases where integration extends to maximum gas thickness 1, 2 and 4 bremsstrahlung radiation lengths. Magnetic pair production⁵ causes the drop at 10^{16} eV. The dip starting at 10^{14} eV is due to pair production on primeval photons, assuming 12 kpc distance. The dotted line represents a non-magnetic cascade, initiated by 5×10^{16} eV protons (integrated to a maximum thickness of 16 radiation lengths). Data: Δ , ref. 19; $\blacktriangle$, ref. 20; $\diamond$, ref. 17; $\blacksquare$, ref. 18; ∇ , ref. 21; $\blacktriangledown$, ref. 22; $\square$, ref. 23; $\circ$, ref. 3; $\bullet$, ref. 4.

of protons from their regular streaming when this regular guiding field is distorted on meeting the gas.

As the pulsar enters eclipse, we must see it through successively greater thicknesses of gas, and hence should see an electron-photon cascade emerging. Figure 1 shows the time-averaged photon spectrum generated by a monoenergetic 10^{17} eV proton beam, assuming that the duration of viewing through different column densities μ of gas varies as μ^{-1} (as in an exponential atmosphere), although this is not critical. The density is typically $\sim 4 \times 10^{-9}$ g cm⁻³, and if there is an embedded field > 10 G, the cascade develops by alternate pair production and very rapid synchrotron radiation rather than by normal bremsstrahlung (and there is little contribution from column densities exceeding 2–4 conventional bremsstrahlung radiation lengths of ~ 40 g cm⁻²). A photon of energy E_γ eV is typically radiated by an electron of energy $E_e = 5.5 \times 10^9 (E_\gamma/B)^{1/2}$ eV (B in gauss). (Alternatively, without a local magnetic field, normal bremsstrahlung dominates: the ultimate spectrum and energy requirements are virtually the same, but more radiation lengths are required, partly because E_γ then falls as E_e rather than E_e^2 . The dotted line in Fig. 1 shows how the photon spectrum then terminates, for monoenergetic 5×10^{16} eV protons.) By choosing the flux of 10^{17} eV protons to fit the photon flux⁴ around 10^{15} eV, it accounts for the 10^{12} eV photons also: there is little room for additional production by protons below 10^{16} eV. With 10^{17} eV protons and a synchrotron cascade, there is no collimation problem. The 10^{12} eV photon is radiated by a 2×10^{14} eV electron (if $B = 500$ G) which suffers negligible deflection in travelling its typical radiation mean free path $10^5/(E_{14}B^2)$ cm, and the direction of the original 10^{17} eV proton is maintained. (There will, nevertheless, be a few widely spread photons, arising from muons which decay after many magnetic gyrations.)

The accelerated beam need not be strictly monoenergetic; we require the energy carried by protons above 5×10^{16} eV to be greater than that carried by protons in the four decades of energy below this, which can also contribute to the TeV photons, and the beam energy could be higher than 10^{17} eV, especially if hadron interactions greatly change⁸ at energies beyond the range explored by accelerators and fitted by the interaction model. Nuclei of the same energy per nucleon are permitted.

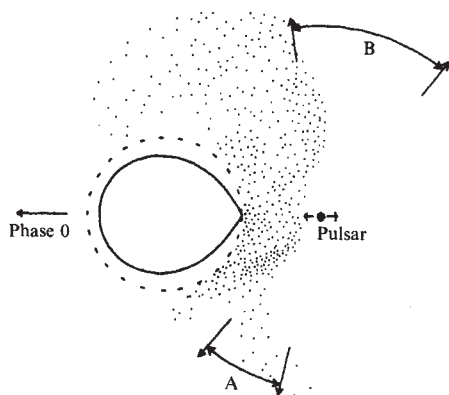


Fig. 2 Binary system, showing approximate directions to the observer when γ -ray bursts are observed (A, phase ~ 0.2 ; B, phase ~ 0.75), and fountain of ejected matter.

To give the flux shown in Fig. 1, the energy in the proton beam above 10^{16} eV must be $2 \times 10^{39} (\Omega/4\pi)(0.05/\Delta) (D/12 \text{ kpc})^2 \text{ erg s}^{-1}$, where D is the distance to Cygnus X-3, Ω is the solid angle into which the overall proton beam is emitted, and Δ is the duty cycle of one photon pulse at high energies (which may be as small as 0.02 (refs 3, 4)). Thus, the power L_p in the particle beam is about $10^{39} \text{ erg s}^{-1}$. Most of the protons (or nuclei) will escape into the Galaxy, providing $\sim 10^{39} \text{ erg s}^{-1}$ above 10^{16} eV per nucleon, and little of significance below this energy.

Wdowczyk and Wolfendale⁹ have already pointed out that about 30 Cygnus X-3s might supply the power to maintain the Galaxy's cosmic rays; but in the energy range above 10^{16} eV, one finds a more remarkable situation. How much power is required to maintain the galactic cosmic rays above 10^{16} eV against leakage? A transport model¹⁰ indicates a lifetime in the galactic disk (say 15 kpc radius and 0.6 kpc thick) of around 2.5×10^5 yr at this energy, which is also compatible with the lifetime of $> 10^6$ yr at 10^{14} eV suggested by Kiraly *et al.*¹¹. Their local energy density $\sim 3 \times 10^{-17} \text{ erg cm}^{-3}$ would then require a power of $\sim 5 \times 10^{37} \text{ erg s}^{-1}$. Alternatively, one can use the variation of anisotropy with energy to estimate¹² the energy spectrum emerging from galactic sources to be $\sim E^{-2.47} dE$ particles s^{-1} (rather like the lower-energy source spectrum $E^{-2.44} dE$ of heavier nuclei deduced from spallation¹³). In this case, $\sim 5 \times 10^{-4}$ of all injected cosmic ray energy goes into the particles above 10^{16} eV, that is, $2.5 \times 10^{37} \text{ erg s}^{-1}$. Hence, the Galaxy's needs are evidently $\leq 10^{38} \text{ erg s}^{-1}$ averaged over 10^5 yr, and sources like Cygnus X-3 can be present only for a part of this time interval. It is therefore unlikely that any other such objects exist in the Galaxy. (The source Vela X-1 is only $\sim 10^{-2}$ as powerful¹⁴.) If the source emits a monoenergetic spectrum, it is normally assumed that one must rely on evolution of beam energy with time to generate the observed spectrum of galactic cosmic rays¹⁵. Cygnus X-3 might, of course, be less powerful than assumed if its pulsar is an aligned rotator, only firing its protons towards us (small solid angle Ω).

Where it strikes the star directly, the beam will deposit $\sim 40\%$ of its energy in an electron-photon cascade at depths 100–500 g cm^{-2} . Here, at $z \sim 100$ Thomson scattering mean free paths below the surface, the thermal leakage time is long, and it will generate a pressure $0.5 L_p z / 4\pi d^2 c$ (if the absorbing layer is at a distance d from the pulsar), z times the direct pressure of the beam. This will easily lift off the overlying few hundred g cm^{-2} if $M_{\text{star}} \sim (0.3\text{--}3) M_{\odot}$, and should be far more effective than keV X-ray absorption in ejecting matter. (The mechanical work done in removing $2 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ —taken as an upper limit⁶—is $< 10^{38} \text{ erg s}^{-1}$.) Nearby surface matter must therefore be lifted off, and a fountain of gas under the pulsar may normally cushion the main body of the star from catastrophic ejection. Although the gas dynamics will be complex¹⁶, a rough idea of the possible disposition of gas is suggested in Fig. 2, based on trajectories of independent particles emitted at up to the escape velocity,

and influenced by a slightly super-eddingtonian radiation pressure from the pulsar. Rotation of the system is responsible for the asymmetry, and the pattern roughly reflects the actual direction from which γ -ray pulses are reported. The large size of the occulting body may arise in this way. Dowthwaite *et al.*¹⁷, however, report a 10^{12} eV signal at phase as early as 0.63, when the pulsar is presumably well in front of the companion. A small X-ray spike is often seen at this point, and as the pulsar is probably plunging in towards periastron near here⁷, it may become embroiled with the gas at this moment. The 34-day variation may be related to a lag in rotation of the companion, along with its magnetic dipole, as this may affect the gas flow.

A requirement of this model is that the 10^{15} eV γ -ray burst should be slightly offset in time (by about half the burst duration—say 0.02 cycle) from the 10^{12} eV γ rays, which require more matter for their production (TeV photons earlier at the 0.2 phase). This may be the case^{4,18}, but there has been uncertainty in the synchronization of phases between different workers, observing over different periods, and it would be desirable to test this by measuring the relative timing of pulses of high and lower energy over the same interval. The 0.7-phase burst may fluctuate somewhat in position, and if so, the start of the sharp intensity drop in X rays should remain synchronized with this γ -ray burst.

Received 5 June; accepted 28 August 1984.

1. Vestrand, W. T. & Eichler, D. *Astrophys. J.* **261**, 251–258 (1982).
2. Eichler, D. & Vestrand, W. T. *Nature* **307**, 613–614 (1984).
3. Samorski, M. & Stamm, W. *Astrophys. J. Lett.* **268**, L17–L21 (1983).
4. Lloyd-Evans, J. *et al. Nature* **305**, 784–787 (1983).
5. Stephens, S. A. & Verma, R. P. *Nature* **308**, 828–830 (1984).
6. Vestrand, W. T. *Astrophys. J.* **271**, 304–314 (1983).
7. Ghosh, P., Elsner, R. F., Weisskopf, M. C. & Sutherland, P. G. *Astrophys. J.* **251**, 230–245 (1981).
8. Wdowczyk, J. & Wolfendale, A. W. *Nature* **306**, 347–349 (1983).
9. Wdowczyk, J. & Wolfendale, A. W. *Nature* **305**, 609–610 (1983).
10. Berezhinsky, V. S., Mikhailov, A. A. & Syrovatskii, S. I. *Proc. 16th Int. Conf. Cosmic Rays*, Kyoto, Vol. 2, 86–91 (University of Tokyo, 1979).
11. Kiraly, P., Kota, J., Osborne, J. L., Stapley, N. R. & Wolfendale, A. W. *Riv. nuovo Cim.* (Ser. 3) **2**, No. 7 (1979).
12. Hillas, A. M. *A. Rev. Astr. Astrophys.* **22**, 425–444 (1984).
13. Koch-Miramond, L. *19th Rencontre de Moriond: 4th Moriond Astrophys. Meet.* (Frontières, Dreux, in the press).
14. Protheroe, R. J., Clay, R. W. & Gerhardt, P. R. *Astrophys. J.* **280**, L47–L50 (1984).
15. Gunn, J. E. & Ostriker, J. P. *Phys. Rev. Lett.* **22**, 728–731 (1969).
16. Bignami, G. F., Marschi, L. & Treves, A. *Astr. Astrophys.* **55**, 155–158 (1977).
17. Dowthwaite, J. C. *et al. Astr. Astrophys.* **126**, 1–6 (1983).
18. Neshpor, Yu. I. *et al. Astrophys. Space Sci.* **61**, 349–355 (1979).
19. Lamb, R. C., Godfrey, C. P., Wheaton, W. A. & Turner, T. *Nature* **296**, 543–544 (1982).
20. Stepanian, A. A., Fomin, V. P., Neshpor, Yu. I., Vladimirov, B. M. & Zyskin, Yu. I. *Proc. Int. Workshop on Very High Energy Gamma-Ray Astronomy*, Ootacamund (eds Ramana Murthy, P. V. & Weekes, T. C.) 43–63 (Tata Institute of Fundamental Research and Smithsonian Institution, Bombay, 1982).
21. Danaher, S., Fegan, D. J., Porter, N. A. & Weekes, T. C. *Nature* **289**, 568–569 (1981).
22. Mukanov, J. B., Nesterova, N. M., Stepanian, A. A. & Fomin, V. P. *Proc. 16th Int. Conf. Cosmic Rays*, Kyoto, Vol. 1, 143–146 (University of Tokyo, 1979).
23. Morello, C., Navarra, G. & Vernetto, S. *Proc. 18th Int. Conf. Cosmic Rays*, Bangalore, Vol. 1, 127–130 (Tata Institute of Fundamental Research, Bombay, 1983).

Rising atmospheric carbon dioxide concentrations may increase streamflow

S. B. Idso* & A. J. Brazel†

*US Water Conservation Laboratory, 4331 East Broadway, Phoenix, Arizona 85040, USA

†Laboratory of Climatology, Arizona State University, Tempe, Arizona 85287, USA

Historically, studies of the greenhouse effect of carbon dioxide (CO_2) have dealt primarily with temperature and only secondarily with precipitation. In the latest report on this topic¹, however, the subject of streamflow is broached with an analysis² which suggests that watersheds in the western United States will suffer 40–75% reductions in streamflow for a doubling of the atmospheric CO_2 content, leading to a 2°C rise in air temperature and a 10% drop in precipitation. A shortcoming of that study is that it does not include the direct antitranspirant effect of atmospheric CO_2 enrichment that would accompany any CO_2 -induced climatic change, whereby increasing CO_2 content of the air tends to induce partial stomatal closure, so reducing plant transpiration and thereby

conserving soil moisture and increasing runoff to streams. Inclusion of this latter effect in a simple model of watershed runoff applied to 12 drainage basins in Arizona indicates that 40–60% increases in streamflow may well be the more likely consequences of a CO₂ concentration doubling, even in the face of adverse changes in temperature and precipitation.

To ensure valid comparability with the National Research Council (NRC) study¹, we used the identical model employed by them to estimate changes in runoff induced by changes in temperature and precipitation and then superimposed the anti-transpirant effect of atmospheric CO₂ enrichment on those results. Thus, we will first briefly describe that model. Table 1 represents a set of empirical relationships among temperature, precipitation and runoff derived by Langbein *et al.*³ for relatively arid areas in the western United States. Our first step was to confirm that this tri-factor matrix used by Revelle and Waggoner² did indeed apply to the 12 Arizona drainage basins we selected for study, which are described in detail elsewhere⁴. Thus, we first derived mean annual values of surface air temperature and precipitation from a report⁵ which gave this information at grid points distributed at spacings of 16 km, and used the results to calculate watershed runoff by this method. Then, from another report⁶ which gave the long-term (30–50 yr) mean annual runoff of the 12 drainage basins as directly measured by stream gauges, we extracted data to use as a standard for comparison. This exercise yielded the plot of Fig. 1, where it can be seen that the Langbein model does indeed adequately describe the mean annual runoff of the 12 Arizona drainage basins selected for study.

We next used the Langbein model of Table 1 in a manner analogous to that of the NRC study to determine the effects of possible CO₂-induced changes in climate. The scenarios we considered were: first, the same 2 °C rise in temperature and 10% drop in precipitation investigated by the NRC report authors; second, a 2 °C rise in temperature with no change in precipitation; and third, a 10% drop in precipitation with no change in temperature. Mean percentage decreases in streamflow for the five wettest basins derived for these scenarios were, respectively, 41, 23 and 23%, with some of the drier basins actually suffering 100% reductions and ceasing to provide any streamflow whatsoever. Up to this point all the results accord with those of Revelle and Waggoner².

Following the lead of the NRC report, we next expressed

Table 1 Mean annual runoff (mm yr⁻¹) as a function of mean annual temperature (°C) and mean annual precipitation (mm yr⁻¹) for the arid southwestern United States

Temperature	Precipitation					
	200	300	400	500	600	700
-2	54	92	154	230	330	440
0	40	74	124	190	275	380
2	28	57	95	154	225	320
4	17	40	78	125	190	265
6	9	25	60	100	155	220
8	0	17	42	82	128	185
10		8	29	64	103	155
12		0	19	47	80	130
14			10	32	65	105
16			0	20	50	85

Sources: refs 2, 3.

runoff (R) as the difference between precipitation (P) and evapotranspiration (E), setting $R = P - E$. Knowing R and P , we could thus solve this equation for E for current conditions, as well as for the three climate change scenarios previously described, and for each of the 12 drainage basins. Then, making use of the experimentally-documented fact that plant evaporative water loss is generally reduced to about two-thirds of its original rate by a doubling of the atmospheric CO₂ content^{7,8}, we adjusted each of these evapotranspiration results to reflect what they would be with the antitranspirant effect of atmospheric CO₂ enrichment superimposed on them. In doing this, however, it was necessary to express the adjusted evapotranspiration (E_a) as $(1 - 1/3 f)E$, where f represents the fractional area of the drainage basin covered by vegetation, as none of the basins was completely covered with plants. Thus, if there were no vegetation ($f = 0$) there would be no vegetative effect and E_a would equal E , whereas if the entire watershed were covered with vegetation, ($f = 1$), E_a would equal $2/3 E$. Our evaluations of f for this purpose were made from detailed descriptions and photographs of the 12 basins associated with the most recent vegetation maps of the area⁹. Values of f obtained by this method varied from a low of 0.24 to a high of 0.61, with an average for the five wettest basins of 0.49.

Following this procedure to obtain E_a values for each of the

Table 2 Mean annual observations and calculations of runoff for the five wettest basins studied for several climate change scenarios

Basin no.	Observations			Calculated runoff							
	Vegetation fraction	Mean annual T (°C)	Mean annual P (mm)	Mean annual R (mm)	Without vegetation effect			With vegetation effect			
					+2°-10%	+2°-0%	+0°-10%	+2°-10%	+2°-0%	+0°-10%	+0°-0%
4906	0.61	9.3	627	111	78	101	105	177	203	198	215
4940	0.57	9.3	607	105	69	91	97	160	188	183	200
4985	0.42	11.3	559	70	46	56	53	109	126	115	138
4990	0.41	12.6	588	74	38	49	58	105	129	122	143
5053	0.45	11.9	562	108	43	62	51	112	128	119	176
Mean	0.49	10.9	589	93	55	72	72	132	155	147	174

Basin no.	% Change from observed							
	Without vegetation effect			With vegetation effect				
	+2°-10%	+2°-0%	+0°-10%	+2°-10%	+2°-0%	+0°-10%	+0°-0%	
4906	-30	-9	-5	+59	+83	+78	+94	
4940	-34	-13	-8	+52	+79	+74	+90	
4985	-34	-20	-24	+56	+80	+64	+97	
4990	-49	-34	-22	+42	+74	+65	+93	
5053	-60	-43	-53	+4	+19	+10	+63	
Mean	-41	-23	-23	+42	+67	+58	+87	

+2°-10% = 2 °C rise in temperature with a 10% reduction in precipitation.

+2°-0% = 2 °C rise in temperature with no change in precipitation.

+0°-10% = no change in temperature with a 10% reduction in precipitation.

+0°-0% = no change in either temperature or precipitation.

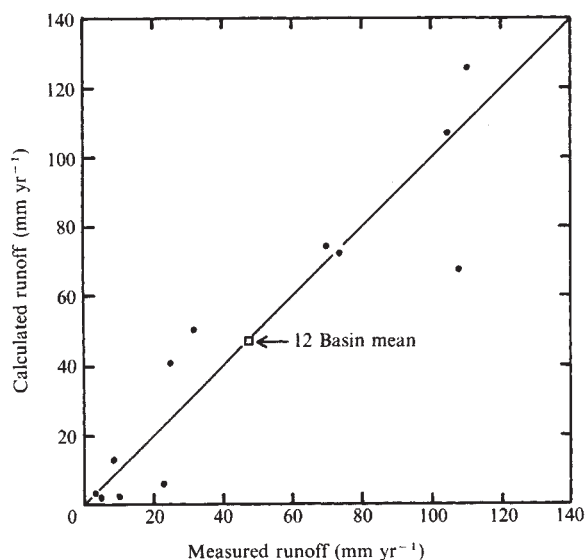


Fig. 1 Mean annual runoff from 12 Arizona drainage basins calculated by the Langbein model of Table 1 compared with measured mean annual runoff.

12 basins for the four scenarios under consideration (three climate change scenarios and one no-change situation), we next used these results to evaluate adjusted runoff (R_a) values which incorporated both the effects of the selected climatic changes (or non-change) plus the direct antitranspirant effects of a CO_2 concentration doubling. The results of this process, summarized in Table 2, indicated that for the first climate change scenario, that is, the 2°C rise in temperature and 10% drop in precipitation predicted in the NRC study, the mean change in streamflow of the five wettest basins calculated to accompany a doubling of the atmospheric CO_2 content went from -41% to $+42\%$ when the antitranspirant effect of the atmospheric CO_2 enrichment was added to the climate change effect. Similarly, when the vegetative effect of CO_2 was added to the climate change effect of the second scenario—a 2°C rise in air temperature with no change in precipitation—the mean change in streamflow of the five wettest basins went from -23% to $+67\%$. And for the third scenario—a 10% drop in precipitation with no change in temperature—it went from -23% to $+58\%$; while for the case of a negligible CO_2 effect on climate, it went from the status quo to $+87\%$. In all of these situations, we note that streamflow increases for the drier basins were of the order of several hundred percent.

We note also that our results are comparable with those recently obtained by Aston¹⁰ with a distributed deterministic process model used to simulate the effects of changed stomatal resistance on streamflow in Australia. Indeed, Aston concluded that “we can expect streamflow to increase from 40 to 90% as a consequence of doubling the atmospheric CO_2 concentration”, in what could also be a statement of our conclusions.

There are several simplifications and uncertainties associated with the above analysis which preclude its being accepted as the final word on the subject. For one thing, the fractional area of vegetation (f) is difficult to assess very accurately. In addition, it could well increase significantly in a CO_2 -enriched world¹¹, and it is likely that an unvegetated area will provide more runoff than a vegetated area of equal size. Also, a small infiltration component of the total water balance is neglected in both our analysis and that of Revelle and Waggoner². Both studies additionally suffer from lack of monthly resolution, and both fail to account for shifts in the time of major watershed runoff which may (or may not) occur as climate changes (or remains the same¹²). However, because both studies have about the same degree of complexity, the very opposite results which they produce illustrate that we are still at the stage where plausible additions to state-of-the-art model studies of the CO_2 ‘problem’

can dramatically alter current preceptions of future consequences.

Received 24 May; accepted 4 September 1984.

1. Carbon Dioxide Assessment Committee, US National Research Council *Changing Climate* (National Academy Press, Washington DC, 1983).
2. Revelle, R. R. & Waggoner, P. E. in *Changing Climate* 419–432 (National Academy Press, Washington DC, 1983).
3. Langbein, W. B. *et al.* *Annual Runoff in the United States*, U.S. Geol. Surv. Circ. No. 5 (US Department of the Interior, Washington, DC, 1949).
4. Brazel, A. J. & Idso, S. B. *Implications of the Rapidly Rising CO_2 Content of Earth's Atmosphere for Water Resources in Arizona* (Climatol. Publ. Sci. Pap. No. 19, Laboratory of Climatology, Arizona State University, 1984).
5. Brazel, A. J. *et al.* *Final Report Arizona Climate Inventory Vol 1* (Laboratory of Climatology, Arizona State University, 1981).
6. US Geological Survey *Water Resources Data Arizona Water Year 1981* (Water-Data Report AZ-81-1, 1983).
7. Kimball, B. A. & Idso, S. B. *Agric. Wat. Managmnt* 7, 55–73 (1983).
8. Rogers, H. H., Thomas, J. F. & Bingham, G. E. *Science* 220, 428–429 (1983).
9. Brown, D. E. (ed.) *Biotic Communities of the American Southwest—United States and Mexico*, *J. Desert Plants* Vol. 4, No. 1–4 (Boyce Thompson Arboretum, 1982).
10. Aston, A. R. *J. Hydrol.* 67, 273–280 (1984).
11. Idso, S. B. & Quinn, J. A. *Vegetational redistribution in Arizona and New Mexico in Response to a Doubling of the Atmospheric CO_2 Concentration* (Climatol. Publ. Sci. Pap. No. 17, Laboratory of Climatology, Arizona State University, 1983).
12. Idso, S. B. *Carbon Dioxide: Friend or Foe? An Inquiry into the Climatic and Agricultural Consequences of the Rapidly Rising CO_2 Content of Earth's Atmosphere* (IBR Press, Tempe, 1982).

Nd evidence for Proterozoic crustal development in the Belt–Purcell Supergroup

C. D. Frost* & R. K. O’Nions

Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, UK

A knowledge of the chemical and isotopic characteristics of the continents as they existed in the past is essential to an understanding of their evolution^{1–7}. Vestiges of these characteristics remain in the sedimentary record. For example, the character and crustal residence age⁴ of ancient continental crust has been deduced from the Sm–Nd isotopic patterns of clastic sediments that have survived to the present day^{4,6,8}. We report here Sm–Nd isotopic measurements from the Belt–Purcell Supergroup of western North America which form the first detailed isotopic investigation of a major Proterozoic sedimentary succession. The Belt–Purcell has many features also seen in the detrital load of the world’s major contemporary rivers, including exceptionally homogeneous $^{147}\text{Sm}/^{144}\text{Nd}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios and a small range of crustal residence ages. Mantle additions to the continental crust seem to have been relatively minor since the middle Proterozoic.

The Belt–Purcell Supergroup of the northwestern United States and southwestern Canada is the most extensive and least disturbed succession of Proterozoic sediments in North America. It outcrops over an area greater than $125,000\text{ km}^2$ in western Montana, northern Idaho, southwestern Alberta and southern British Columbia (Fig. 1), and lies on $\geq 1.7\text{ Gyr}$ crystalline basement^{9,10}. Shallow water structures predominate throughout the succession^{11–15}. The time of deposition of the supergroup is uncertain, but may have spanned an interval of 1,500–1,200 Myr^{10,16–18}. Unconformably overlying the Belt–Purcell Supergroup are the impure clastics of the Windermere Supergroup, deposited around 800–600 Myr^{15,19}.

Proposed explanations for the depositional environment of the Belt–Purcell Supergroup include deposition in a miogeocline or within an adjacent failed arm of a rift system, and suggestions that Belt–Purcell deposition coincided with incipient rifting of a western continent, of entirely epicratonic deposition^{15,20–24}. Sediment source areas have been tentatively identified to the east, south and south-west of the Belt–Purcell deposit^{12,15,25,26}. However, Mesozoic listric thrust faulting within the Belt–Purcell basin involved translations of up to 180 km ¹¹ and thus hampers attempts to elucidate the original basin configuration, sediment

* Present address: Department of Geology and Geophysics, University of Wyoming, Laramie, Wyoming 82071, USA.

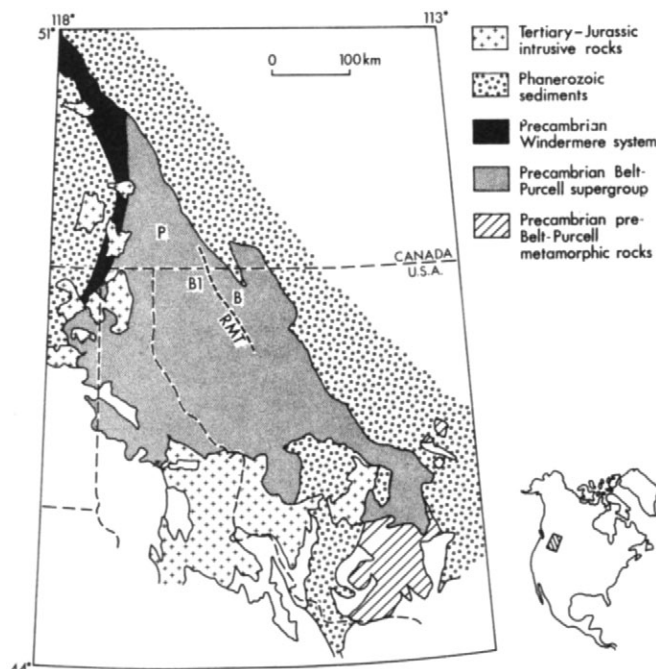


Fig. 1 Geological sketch map showing the location of the Belt-Purcell Supergroup, principal structures, and sample localities. Modified from King³⁴. RMT, Rocky Mountain Trench; B, B1, Belt Series; P, Purcell Series.

source directions or source areas. Information from Sm-Nd isotopes obtained from Belt-Purcell sediments may provide new insights into these problems, as well as into Proterozoic crustal development.

Samples have been selected mainly from two localities, one in the Purcell Mountains, British Columbia, and the other from the vicinity of the Whitefish Range, Montana (Fig. 1). We also analysed a sample from the Horsethief Creek Group of the Windermere Supergroup. Exact sample locations and stratigraphic positions are described elsewhere²⁷.

Table 1 gives Sm-Nd isotopic measurements on these samples. Analytical procedures used follow those reported previously²⁸. Present-day $^{143}\text{Nd}/^{144}\text{Nd}$ ratios for Belt-Purcell rocks fall within a restricted range of 0.511615 ± 18 to 0.512040 ± 22 ($\epsilon_{\text{Nd}} = -20$ to -12) and $^{147}\text{Sm}/^{144}\text{Nd}$ ratios vary only from 0.099 to 0.120, even though Nd and Sm concentrations vary by a factor of 3–4. These $^{147}\text{Sm}/^{144}\text{Nd}$ ratios are typical of those reported for modern river sediments²⁹ and other fine-grained clastic sedimentary rock^{5,6,8}.

The uniformity of $^{143}\text{Nd}/^{144}\text{Nd}$ ratios for sediments deposited throughout a ~20-km thick succession during a period that may extend over some 300 Myr is striking (Fig. 2a). This constant isotope ratio probably indicates the presence of a particularly homogeneous and persistent sediment source, and/or very thorough mixing of a number of source materials over a period of several hundred million years. Crustal residence ages, which are calculated on the assumption that the Sm/Nd ratios of the sediments are unchanged from the time that their precursory continental material was generated by magmatic processes, are also uniform for the Belt-Purcell sediments, and with one exception fall between 1.9 and 2.1 Gyr (Table 1, Fig. 2b). The mean crustal residence age of the uniform source of sediment is well in excess of the depositional age and new mantle additions to the continent within the Belt-Purcell source area, over the interval concerned, are too small to be recognizable. The one sample which possesses a somewhat younger crustal residence age of 1.6 Gyr presumably reflects a larger than average additional contribution from a relatively young and probably locally supplied source.

It is useful to compare crustal residence and depositional ages of Belt-Purcell sediments with published results for Archaean sediments. The crustal residence age for Belt-Purcell samples

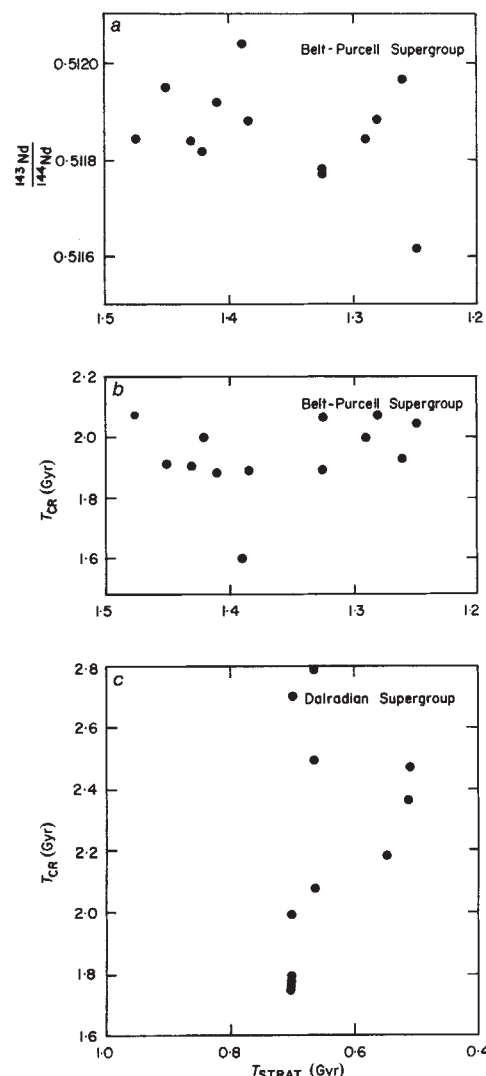


Fig. 2 a, $^{143}\text{Nd}/^{144}\text{Nd}$ plotted with respect to approximate stratigraphic ages for Belt-Purcell sediments. Stratigraphic ages are inferred from palaeomagnetic data¹⁶. b, Crustal residence age estimates (T_{CR}) plotted with respect to approximate stratigraphic age (T_{STRAT}) for Belt-Purcell sediments. c, Crustal residence age estimates (T_{CR}) plotted with respect to approximate stratigraphic age (T_{STRAT}) for Dairadian Supergroup samples. Comparison of b and c illustrates that whereas the Belt-Purcell shows a small range in crustal residence age throughout a 20-km succession, the Dairadian shows an unusually wide variation in crustal residence age. (See also ref. 35.)

exceeds the depositional age by 200–800 Myr, whereas available Sm-Nd analyses of Archaean sediments from Isua, Malene, Pilbara and Fig Tree deposits indicate crustal residence ages approximately coincident with their stratigraphic ages^{4,6,8,28}. This result has been interpreted to indicate that these Archaean sediments are close to first-cycle⁴, because short time intervals around 100 Myr have elapsed between extraction of parent material from the mantle and deposition of the sediment. The contrasting character of the Archaean and Belt-Purcell sediments indicates that the latter has a longer average time interval between crustal formation and sediment deposition and possibly a more extensive history of recycling. The Windermere Supergroup sample (Table 1) has a lower $^{143}\text{Nd}/^{144}\text{Nd}$ ratio (0.51154) than that of the Belt-Purcell samples, and higher crustal residence age of 2.4 Gyr, perhaps reflecting the contribution of a different, older source material or proportion of such material to the Windermere.

The sedimentology of the Belt-Purcell rocks does not constrain their source. However, to evaluate one of several possible source areas, eight Canadian Shield basement samples have

Table 1 Nd and Sm isotope data for fine-grained clastic sediments from the locations within the Belt–Purcell Supergroup

Belt Series: Whitefish Range and vicinity						
Sample	$^{143}\text{Nd}/^{144}\text{Nd}(2\sigma)^*$	Nd (p.p.m.) †	Sm (p.p.m.) †	$^{147}\text{Sm}/^{144}\text{Nd}$	T_{CR} (Gyr)	T_{STRAT} (Gyr)
B1	0.511842 (20)	24.78	4.814	0.1174	2.07	1.47
B2	0.511836 (18)	36.13	6.639	0.1111	1.95	1.43
B3	0.511917 (16)	44.22	8.329	0.1139	1.88	1.41
B4	0.512040 (22)	41.19	7.396	0.1072	1.59	1.39
B5	0.511771 (20)	47.41	8.799	0.1122	2.07	0.32
B6	0.511965 (14)	28.45	5.665	0.1204	1.93	1.26
Purcell Series: Purcell Mountains and vicinity						
P1	0.511950 (18)	42.21	8.238	0.1180	1.91	1.45
P2	0.511814 (18)	44.86	8.219	0.1107	1.97	1.42
P3	0.511879 (18)	32.16	5.891	0.1107	1.88	1.38
P4	0.511764 (15)	38.26	6.460	0.1021	1.89	1.32
P5	0.511842 (18)	52.93	10.04	0.1146	2.01	1.29
P6	0.511882 (24)	58.98	11.76	0.1205	2.07	1.28
P7	0.511615 (18)	20.40	3.348	0.0992	2.04	1.25
Windermere Supergroup: Purcell Mountains and vicinity						
P8	0.511537 (20)	36.04	6.543	0.1097	2.63	<0.750

$^{143}\text{Nd}/^{144}\text{Nd}$ ratios are normalized to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$. Bulk Earth parameters used to calculate ϵ_{Nd}^T are $^{143}\text{Nd}/^{144}\text{Nd} = 0.512638$ and $^{147}\text{Sm}/^{144}\text{Nd} = 0.1966$. Crustal residence ages (T_{CR}) are calculated relative to a depleted mantle where $\epsilon_{\text{Nd}}^{4.6} = 0$ and $\epsilon_{\text{Nd}}^0 = +10$ (ref. 29). Samples are numbered B1 to B6 and P1 to P7 from the base to the top of the stratigraphic succession collected from the Belt and Purcell Series respectively. T_{STRAT} ages are approximate.

* Errors quoted are 2 standard deviations of the mean.

† Error $\leq 0.2\%$.

been obtained from cores in southern British Columbia and Alberta recovered from depths of 2,200–3,700 m. The results for these samples (Table 2) are compared with previous Nd isotopic analyses reported from the Canadian Shield⁸. Crustal residence ages for the eastern and western parts of the shield are similar and range respectively over 2.6–3.2 and 2.8–3.4 Gyr, with the one exception of a Grenville composite which has a crustal residence age of 1.5 Gyr. No part of the Canadian Shield has yet been identified which has a 2.0 Gyr crustal residence age such as might have acted as a single eastern sediment source to the Belt–Purcell, and furthermore, no mixture of these Canadian Shield regions west of the Grenville Province yields an average crustal residence age of 2.0 Gyr. If multiple sources with a range of crustal residence ages are involved, they must have become effectively homogeneous by the time of deposition.

The range in crustal residence ages for both modern river particulates²⁹ and Belt–Purcell sediments suggests that the Belt–Purcell material, like modern river sediment, is the product of

effective homogenization accompanying fluvial processes. Lithofacies evidence of braided stream, alluvial fan and deltaic deposits supports this notion, as does the general uniformity of sediment mineralogy and geochemistry^{12,23,26,30}. By analogy with modern major river systems, the Belt–Purcell succession may be the product of erosion across a large drainage area, and sediment sources may not have been restricted to proximal areas.

It should be emphasized that the particularly restricted range of crustal residence age estimates found for the Belt–Purcell is not a feature of all sedimentary successions. For example, the Dalradian Supergroup of Scotland, for which considerable data are available, exhibits very different characteristics (Fig. 2c)^{4,31}. The Dalradian was deposited with an interval extending over 0.70–0.51 Gyr, and shows a very wide range of crustal residence ages (1.7–2.8 Gyr). These results are consistent with localized and heterogeneous source regions, accessed by drainage systems of restricted areal extent. The Dalradian results provide an

Table 2 Nd and Sm isotope data for Canadian Shield samples

New data from basement cores, Southern British Columbia and Alberta						
Sample	$^{143}\text{Nd}/^{144}\text{Nd}(2\sigma)^*$	Nd (p.p.m.) †	Sm (p.p.m.) †	$^{147}\text{Sm}/^{144}\text{Nd}$	T_{CR} (Gyr)	
UA4221	0.511275 (26)	4.125	0.7736	0.1133	2.83	
UA4229	0.511001 (18)	59.02	8.702	0.0891	2.62	
UA4418B	0.511135 (20)	18.37	3.263	0.1074	2.87	
UA4220	0.511448 (24)	21.64	4.065	0.1135	2.58	
UA2438	0.510381 (14)	49.51	4.955	0.0610	2.75	
	0.510363 (14)					
UA4231D	0.511258 (17)	14.31	2.934	0.1239	3.19	
UA3185	0.511217 (18)	47.75	8.615	0.1091	2.80	
UA5282	0.510979 (16)	32.92	5.351	0.0982	2.85	
Previously reported data for Canadian Shield composites ⁸						
Sample			$\epsilon_{\text{Nd}}^{(0)}$	$f_{\text{Sm}/\text{Nd}}^{\ddagger}$	T_{CR} (Gyr) §	
New Quebec			-29.4 ± 0.4	-0.444	2.94	
North Quebec			-28.9 ± 0.3	-0.464	2.79	
Fort Enterprise gneiss			-24.3 ± 0.4	-0.372	3.37	
Fort Enterprise granite			-30.7 ± 0.2	-0.495	2.77	
Baffin Island			-31.7 ± 0.3	-0.464	3.00	
Saskatchewan			-32.5 ± 0.4	-0.486	2.94	
Quebec			-7.1 ± 0.3	-0.361	1.52	

* Errors quoted are 2 standard deviations of the mean.

† Error $\leq 0.2\%$.

$^{\ddagger} f_{\text{Sm}/\text{Nd}} = \frac{(\text{Sm}/\text{Nd})_{\text{sample}}}{(\text{Sm}/\text{Nd})_{\text{CHUR}}} - 1$.

$^{\S} T_{\text{CR}}$ calculated from ϵ_{Nd} and $f_{\text{Sm}/\text{Nd}}$ data of McCulloch and Wasserburg⁸.

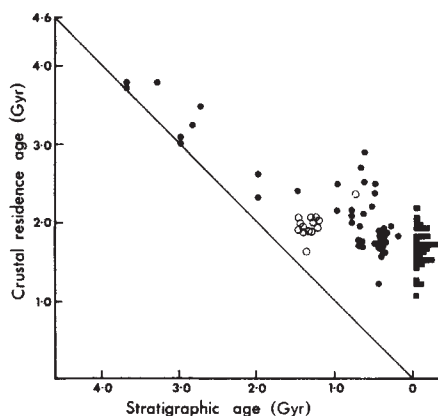


Fig. 3 Comparison of the stratigraphic ages and crustal residence ages of fine-grained clastic sediments calculated from Sm-Nd isotopic analyses from this study and elsewhere^{4,6,8,29,31}. Additions to the sedimentary mass from contemporary mantle will plot near the line indicated. Open circles, Belt-Purcell samples; closed circles, all other samples; closed squares indicate river and aeolian particulates.

excellent example of the value of Sm-Nd isotopic analyses of sediments for limiting the nature of sediment source regions.

A comparison of crustal residence ages for the Proterozoic Belt-Purcell Supergroup and sediments of other depositional ages (Fig. 3) shows that all sediments deposited since 2.0 Gyr have crustal residence ages well in excess of their stratigraphic ages. Assuming that the samples analysed are typical of sediment of their respective stratigraphic ages, it appears that the source areas of sediments deposited before ~2.0 Gyr must include a large component of material with a comparatively short crustal residence time. In contrast, the younger sediments are dominated by recycled continental components with relatively long crustal residence times, and minimal additions of contemporary mantle-derived material. Interestingly, the average crustal residence age for modern river particulates, Palaeozoic sediments and the Belt-Purcell Supergroup are all around 1.7–2.0 Gyr. Although it might be argued that this pattern is an artefact arising from limited sampling and that the Belt-Purcell may be atypical of the global sedimentary mass existing at that time, it seems more likely that the input of new mantle material into the sedimentary mass has been minor over the past 1.5 Gyr or so. Because the evolution of the sedimentary mass is likely to be reasonably representative of at least the upper part of the continental crust, it would appear that the evolution of the continents themselves has also been dominated by recycling.

An important conclusion from the evidence now available is that the mean crustal residence age of clastic sediments, and by inference that of the continental crust during much of the Archaean, differed little from that expected for contemporary additions from the mantle at that time. This observation eliminates the possibility of major contributions from pre-4.0 Gyr continental crust, and suggests that the residence time of light rare-earth enriched material at the Earth's surface was short during this interval^{32,33}. The dominance of material with a short crustal residence age evident in the archaean sedimentary mass is reduced markedly throughout the Proterozoic. Recycling and cannibalism of older continental crust has dominated the Proterozoic and Phanerozoic evolution of the sedimentary mass.

We thank Drs R. A. Burwash, J. Harrison, J. Patchett, Z. Peterman and A. G. Smith for critical reviews. This work has been partially supported by an NSF grant to C.D.F., the Royal Society, and the Natural Environment Research Council. Cambridge Department of Earth Sciences Contribution No. 257.

Received 11 May; accepted 16 July 1984.

1. Hurley, P. M. & Rand, R. J. *Science* **164**, 1229–1242 (1969).
2. Moorbath, S. *Phil. Trans. R. Soc. A* **288**, 401–413 (1978).
3. Armstrong, R. L. *Phil. Trans. R. Soc. A* **301**, 443–472 (1981).
4. O'Nions, R. K., Hamilton, P. J. & Hooker, P. J. *Earth planet. Sci. Lett.* **63**, 229–240 (1983).
5. DePaolo, D. J. *Geophys. Res. Lett.* **10**, 705–708 (1983).

6. Allègre, C. J. & Rousseau, D. *Earth planet. Sci. Lett.* **67**, 19–34 (1984).
7. Patchett, J. & Chauvel, C. *Geophys. Res. Lett.* **11**, 151–153 (1984).
8. McCulloch, M. T. & Wasserburg, G. J. *Science* **200**, 1003–1011 (1978).
9. Giletti, B. J. *J. geophys. Res.* **71**, 4029–4036 (1966).
10. Zartman, R. E., Peterman, Z. E., Obradovich, J. D., Gallego, M. D. & Bishop, D. T. *Idaho Bur. Mines Geol. Bull.* **24**, 61–69 (1982).
11. Harrison, J. E., Kleinkopf, M. D. & Wells, J. D. *Geology* **8**, 407–411 (1980).
12. Harrison, J. E. *Bull. geol. Soc. Am.* **83**, 1215–1240 (1972).
13. Smith, A. G. & Barnes, W. C. *Bull. geol. Soc. Am.* **77**, 1399–1426 (1966).
14. Edmunds, F. R. *Belt Symp.* **1**, 230–234 (1973).
15. Gabrielse, H. *Am. J. Sci.* **272**, 521–536 (1972).
16. Ryan, B. D. & Blenkinsop, J. *Can. J. Earth Sci.* **8**, 85–95 (1971).
17. Elston, D. P. & Bressler, S. L. *J. geophys. Res.* **85**, 339–355 (1980).
18. Obradovich, J. D. & Peterman, Z. E. *Can. J. Earth Sci.* **5**, 737–747 (1968).
19. Bond, G. C., Kominz, M. A. & Devlin, W. J. *Nature* **306**, 775–779 (1983).
20. Monger, J. W. H. & Price, R. A. *Can. J. Earth Sci.* **16**, 770–791 (1979).
21. Stewart, J. H. *Geology* **4**, 11–15 (1976).
22. Høy, T. *Montana Bur. Mines Geol. spec. Publ.* **90**, 36–38 (1984).
23. Cressman, E. *Montana Bur. Mines Geol. spec. Publ.* **90**, 8–9 (1984).
24. Sears, J. W. & Price, R. A. *Geology* **6**, 267–270 (1978).
25. Winston, D. *Belt Symp.* **1**, 208–220 (1973).
26. Winston, D. *Can. Soc. Petrol. Geol. Mem.* **5**, 343–359 (1978).
27. Frost, C. D. thesis, Univ. Cambridge (1983).
28. Hamilton, P. J., O'Nions, R. K., Bridgewater, D. & Nutman, A. *Earth planet. Sci. Lett.* **62**, 263–272 (1983).
29. Goldstein, S. L., O'Nions, R. K. & Hamilton, P. J. *Earth planet. Sci. Lett.* (in the press).
30. Horodyski, R. J. *Precamb. Res.* **20**, 391–425 (1983).
31. Frost, C. D. & O'Nions, R. K. *J. Petrol.* (submitted).
32. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *J. geophys. Res.* **84**, 6091–6101 (1979).
33. O'Nions, R. K., Evensen, N. M. & Hamilton, P. J. *Phil. Trans. R. Soc. A* **297**, 479–493 (1980).
34. King, P. B. *Tectonic Map of North America, 1:500,000*. (U.S. Geological Survey, Washington DC, 1969).
35. Miller, R. G. & O'Nions, R. K. *Earth planet. Sci. Lett.* **68**, 459–470 (1984).

Metamorphism of reduced granulites in low-CO₂ vapour-free environment

William Lamb & J. W. Valley

Department of Geology and Geophysics,
University of Wisconsin, Madison, Wisconsin 53706, USA

Pervasive flooding of CO₂ has been proposed as the cause of granulite facies metamorphism that is capable of producing many distinctive characteristics of the deep continental crust: reduced water activity, orthopyroxene-bearing assemblages, depletion of large-ion lithophile (LIL) elements, and dehydration^{1–6}. The calculations presented here of C–O–H fluid composition for conditions of granulite facies metamorphism show that oxygen fugacity (f_{O_2}) estimates from many terranes are sufficiently low that the addition of CO₂-rich fluid causes graphite to precipitate. For values of pressure (P) and temperature (T) common to granulites, and with f_{O_2} slightly below the quartz–fayalite–magnetite buffer (QFM), the addition of CO₂ sufficient to grow 10 vol. % orthopyroxene requires the precipitation of 1.5 vol. % graphite. As 0.1 vol. % graphite is readily recognizable, but is not reported in most low f_{O_2} granulites, these rocks have not been flooded by CO₂, and low f_{H_2O} is probably due to extraction of a magma or recrystallization of an already dry rock^{7,8}. Granulite terranes may thus result from a combination of these three processes and the dominance of any one cannot now be demonstrated on a regional basis.

Calculations in the C–O–H system show that there are six important fluid species (H₂O, CO₂, CH₄, CO, H₂, O₂) which can be related by four independent equations^{9–11}.



At known P and T there are eight variables in this system, the six fluid species, the fluid pressure (P_F) and the activity of graphite (a_C). If a free-fluid phase is present, a condition required by CO₂ infiltration, then $P_F = P_{Lithostatic}$ (P_L) and equation (5) can be assumed.

$$P_L = P_{H_2O} + P_{CO_2} + P_{CH_4} + P_{CO} + P_{H_2} + P_{O_2} \quad (5)$$

By assuming a free-fluid phase there are now five equations and seven unknowns (as P_F has been fixed) and by fixing the value of two variables it is then possible to calculate the values of the

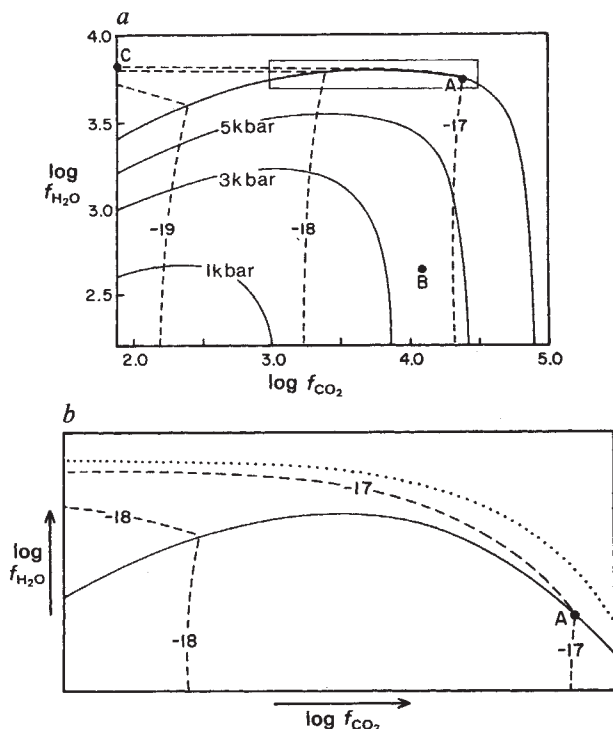


Fig. 1 *a*, *b*, C-O-H fluid equilibria at 700 °C, $P_L = 7$ kbar. Graphite is stable only along the outermost solid curve. Outside this curve $\alpha_C < 1$ at $P_F = 7$ kbar. Inside this curve $P_F < 7$ kbar so that a freely-migrating fluid at lithostatic pressure cannot exist within this region. Isopleths of log f_{O_2} (dashed) show the effect of an oxygen buffering mineral assemblage (these isopleths are calculated assuming $\alpha_C = 1$ within the graphite stability curve). *b*, An enlargement of the area within the box in *a*, showing the intersection of the log $f_{O_2} = -17$ isopleth with the graphite stability curve at point A. At this point $X_{CO_2} \approx 0.3$ and addition of CO₂ will precipitate graphite. For these conditions the amount of CO₂ necessary to stabilize 10 vol. % orthopyroxene would precipitate 1.5 vol. % graphite.

five remaining variables. In a graphite-bearing system where $\alpha_C = 1$, it is thus possible to calculate the fugacity of five fluid species by determining the fugacity of one. If graphite is not present the determination of the fugacity of two fluid species allows the determination of the other four as well as α_C .

The calculations summarized here were performed by the program CALCOH which employs a hard-sphere modified Redlich-Kwong equation of state for H₂O, CO₂ and CH₄ (refs 12, 13), and the Gibbs energy of each fluid¹⁴. Calculations of fluid fugacities were made with and without consideration of non-ideal mixing by these three fluid species, with a general agreement of better than 0.1 log unit.

Results of these calculations are shown as a function of log f_{CO_2} versus log f_{H_2O} at constant $T = 700$ °C, $P_L = 7$ kbar (Fig. 1*a*, *b*). The outermost solid curve delineates the stability field of graphite. Along this curve $\alpha_C = 1$, $P_F = P_L = 7$ kbar, and knowledge of the fugacity of any one of the six fluid species determines the system. Any point inside (below and to the left of) the graphite stability curve is equivalently fixed for a lower P_F as indicated by the 1-, 3- and 5-kbar lines. These pressure contours are calculated assuming $\alpha_C = 1$. It is important to point out that for a graphite-absent rock, where $\alpha_C < 1$, these contours become the upper limit to P_F . Thus, at $P_L = 7$ kbar there will not be a free-fluid phase for values buffered inside of the graphite stability curve as $P_F < P_L$. The dotted curve (Fig. 1*b*) satisfies the relation $P_{H_2O} + P_{CO_2} = 7$ kbar. As high-temperature rocks are ductile with relatively little strength, the region above this curve is geologically inaccessible as a fluid overpressure would result, allowing fluids to escape. The log f_{O_2} contours are calculated at $P_F = 7$ kbar outside (above) the graphite stability curve at $\alpha_C < 1$, and $P_F < 7$ kbar inside graphite stability at $\alpha_C = 1$. Only

on the graphite stability curve can $P_F = 7$ kbar in the presence of graphite ($\alpha_C = 1$). An important ramification of this diagram is the instability of 7 kbar of fluid inside the graphite stability curve.

The significance of this result is seen for a rock where log f_{O_2} is buffered to -17 at 700 °C, 7 kbar (dashed isopleth in Fig. 1). At low f_{CO_2} (near point C) graphite is unstable at $P_F = 7$ kbar. However, if CO₂ infiltrates the rock it becomes the dominant fluid species at a very low CO₂/R as pore space will be low at these P and T and no other fluid species are assumed to be buffered (CO₂/R is the CO₂/rock ratio calculated from the moles of oxygen; some workers prefer to use this weight ratio, the conversion being 1.0 atom ratio $\approx$ 0.5 wt ratio). If infiltrating CO₂ dilutes intergranular H₂O then the system will increase in f_{CO_2} (and α_C) following the log $f_{O_2} = -17$ isopleth towards the right from point C and intersecting the graphite stability curve at point A at which time graphite becomes stable and $X_{CO_2} \approx 0.3$. The system is then invariant in the presence of graphite. Any additional CO₂ infiltration will be buffered and will drive reaction (1) to the left, precipitating graphite and exhausting the oxygen buffering capacity of the rock. As long as this buffering capacity is maintained CO₂ will be converted quantitatively to graphite and CO₂/R for infiltration can be calculated from the amount of graphite that is precipitated. Such CO₂/R values are considered maxima because some graphite might predate and be unrelated to CO₂ infiltration.

An estimate of the amount of CO₂-flooding necessary to form a granulite can be made from modal analysis and consideration of reaction progress for dehydration reactions. For example, based on the idealized stoichiometry of the amphibole breakdown reaction:



it is estimated that in certain rocks from the Adirondack Mountains¹⁵ H₂O was buffered at $X_{H_2O} = 0.12$, at 700 °C and 7 kbar. If 10 vol. % orthopyroxene formed by reaction (6) in response to CO₂ infiltration then CO₂ must have continuously diluted all evolved H₂O. In this example, H₂O must always comprise 12 mol.% of the fluid ($X_{CO_2} \approx 0.88$), and the total CO₂/R is estimated at 0.1. If conditions of $X_{H_2O} \approx 0.1$ from the Adirondacks^{8,15,16} are similar to those of other granulite terranes then values of CO₂/R $\sim$ 0.1 should generally be necessary to form granulites by CO₂ flooding. This value is, however, a conservative estimate and values of CO₂/R = 0.25 have been independently estimated². Higher CO₂/R values would cause precipitation of larger amounts of graphite in reduced conditions.

In a rock with <0.1 vol. % pore space, infiltration equal to CO₂/R = 0.1 will precipitate $\sim$ 1.5 vol. % graphite. At low porosity, the amount of graphite precipitated depends only on CO₂/R and is independent of f_{O_2} as long as f_{O_2} is buffered to within the stability of graphite. If a high pore volume of 1% is chosen then the amount of graphite is negligibly reduced to 1.47%. Regardless of its form, as little as 0.1% graphite is easily identified either macroscopically or in thin section allowing reliable identification of rocks that have undergone very small amounts of infiltration (CO₂/R > 0.007). Thus it would be difficult to miss evidence for CO₂ infiltration in low f_{O_2} rocks.

The general applicability of these calculations is dependent on what values of f_{O_2} are buffered during granulite facies metamorphism. If f_{O_2} is buffered below reaction (1) then large amounts of readily observable graphite must be present to satisfy models of CO₂ flooding, while if f_{O_2} values exceed the stability of graphite then the absence of graphite provides no information as to CO₂/R.

Unfortunately, the presence of graphite by itself does not necessarily indicate CO₂ infiltration even though it obviously signifies f_{O_2} values less than reaction (1). The problem arises because much graphite of premetamorphic, biogenic origin is preserved in granulites¹⁷, and thus the presence of graphite, which may be 2–10 vol. % in some deposits¹⁸, indicates CO₂ infiltration only if lower grade equivalents are graphite-free. Glassley³ concludes that CO₂ flooding is indicated by the

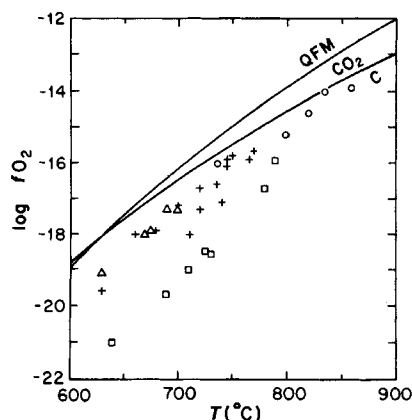


Fig. 2 Values of f_{O_2} for granulite facies samples estimated from coexisting magnetite + ilmenite. Reaction (1), which closely approximates the maximum f_{O_2} stability of graphite, and QFM²⁶, the quartz + magnetite + fayalite O_2 -buffer, are shown for reference at 7 kbar. Samples: ○, from Scourie²⁵; □, from Quebec²⁴; △, from south Quebec²³; +, from the Adirondacks²². Graphite was not reported in most samples indicating that if CO_2 infiltration occurred, amounts of CO_2 were very small, $CO_2/R < 0.007$, and were not sufficient to account for granulite formation.

presence of graphite in two granulite terranes, but this has been criticized on the grounds that graphite-absent amphibolite facies rocks are not correlative with graphite-bearing granulites¹⁹. The great difficulty in correlating across the amphibolite-granulite transition has thus far prevented successful application of this test; however, the presence of small amounts of graphite is still restrictive as it places an upper limit to CO_2 infiltration. For example, it has been estimated that 35% of the rocks in an area of the Marranguit Peninsula, West Greenland contain graphite at levels of $<0.01\%$ (ref. 20). If all of this graphite was precipitated by CO_2 infiltration then a CO_2/R value of <0.0007 is indicated, which is not enough to form these granulites. If some graphite is biogenic in origin then CO_2/R would be less, possibly zero.

Coexisting Fe-Ti oxides have been successfully used to estimate f_{O_2} and temperature in a wide range of igneous and high-grade metamorphic terranes²¹. Values from four studies of granulites are shown in Fig. 2 (Adirondacks²², Quebec^{23,24}, and Scourie²⁵). These values are from several rock types including charnockites, anorthosites, metagabbros, granites, trondhjemites and a titaniferous magnetite deposit. Because most f_{O_2} values plot below reaction (1), even when considering an uncertainty of $\log f_{O_2} \pm 0.5$, it is clear that most of these samples should contain graphite if they were infiltrated by CO_2 . No graphite is reported for any of the samples from refs 22, 23 and 25. We have further examined the Adirondack samples²², both in thin section and macroscopically, but have observed no graphite at the 0.01% level.

The assemblage, biotite + K-feldspar + magnetite + ilmenite, buffered f_{O_2} and f_{H_2O} in an Adirondack charnockite. The estimated values are $\log f_{O_2} = -17.2 \pm 0.4$ and $\log f_{H_2O} = 2.65 \pm 0.65$ (ref. 16). These values are plotted in Fig. 1 (point B), showing that if $a_C = 1$ then P_F must have been ~ 4 kbar for this sample (if $a_C < 1$ then P_F was less). Thus, these estimates indicate that no free-fluid phase could have coexisted with this rock at the pressure and temperature estimated for granulite facies metamorphism, unless the metamorphic fluid deviated substantially from the C-O-H system. We consider the presence of such great amounts of other fluid species (such as H_2S , N_2 , Ar, HF, NaCl) to be unlikely on a regional basis. Common orthopyroxene-bearing assemblages in granulites indicate $\log f_{H_2O}$ is significantly below 3.5 ($X_{H_2O} \leq 0.1$), and such low f_{H_2O} is generally ascribed to all granulite terranes based on melting equilibria as well as dehydration reactions. If the rocks in Fig. 2 were buffered at low f_{H_2O} then none of these samples could have been infiltrated by significant CO_2 as a free C-O-H fluid at $P_F = 7$ kbar would

not be stable. It is thus likely that all of these rocks represent fluid-absent granulite facies metamorphism.

Further studies are necessary before it is possible to generalize about the fluid conditions of the lower crust. These results show that some common, granulite facies rocks have not been flooded by sufficient CO_2 to explain their genesis. However, compelling evidence indicates that granulites from certain quarries in southern India were stabilized by infiltrating CO_2 (refs 2, 4, 6). Graphite is not reported in these rocks, but since f_{O_2} is buffered to above QFM by the reaction $3\text{ferrosilite} + 1/2 O_2 = \text{magnetite} + 3\text{quartz}$ (R. Newton, personal communication), graphite is not expected regardless of whether CO_2 flooding occurred or not. Thus the results reported here may not be at variance with those from southern India. Extreme gradients in f_{CO_2} and the absence of pervasive fluid movement is reported for some localities in the Adirondacks⁸. If such heterogeneous fluid conditions are representative of granulite facies terranes elsewhere, then low f_{H_2O} values may indicate either extraction of magmas, channelized CO_2 infiltration, or metamorphism of already dry rock, and these processes may operate in close proximity. Very real contrasts in process may also exist from terrane to terrane due to differences in age and/or tectonic style.

This study was supported by grants from the Gas Research Institute (5083-260-0852), the NSF (EAR81-21214 and 83-11772) and the Geological Society of America Penrose Fund (3256-83). We thank P. E. Brown, T. C. Juster, E. J. Essene and W. E. Glassley for helpful conversations and reviews.

Note added in proof: We have used the calibration of Buddington and Lindsley²¹ to estimate f_{O_2} from Fe-Ti oxides rather than that of Spencer and Lindsley²⁷ (1981) (D. H. Lindsley, personal communication). More recently, unpublished experiments of Anderson and Lindsley yield somewhat lower T at low f_{O_2} (D. H. Lindsley, personal communication) and suggest that the oxides from Quebec²⁴ have reset after metamorphism. Most of the other values in Fig. 2 are not significantly affected and thus the conclusions of the study are not changed.

Received 11 May; accepted 24 July 1984.

1. Touret, J. *Lithos* **4**, 239-249; 423-436 (1971).
2. Newton, R. C., Smith, J. V. & Windley, B. F. *Nature* **288**, 45-50 (1980).
3. Glassley, W. E. *Nature* **295**, 229-231 (1982).
4. Janardhan, A. S., Newton, R. C. & Hansen, E. C. *Contr. Miner. Petrol.* **79**, 130-149 (1982).
5. Glassley, W. E. *Geochim. cosmochim. Acta* **47**, 596-616 (1983).
6. Condie, K. C., Allen, P. & Narayana, B. L. *Contr. Miner. Petrol.* **81**, 157-167 (1982).
7. Fyfe, W. S. *Phil. Trans. R. Soc. Lond.* **273A**, 457-461 (1973).
8. Valley, J. W. & O'Neill, J. R. *Contr. Miner. Petrol.* **85**, 158-173 (1984).
9. French, B. M. *Rev. Geophys.* **4**, 223-253 (1966).
10. Ohmoto, H. & Kerrick, D. M. *Am. J. Sci.* **277**, 1013-1044 (1977).
11. Valley, J. W., Peterson, E. U., Essene, E. J. & Bowman, J. R. *Am. Miner.* **67**, 545-557 (1982).
12. Kerrick, D. M. & Jacobs, G. K. *Am. J. Sci.* **281**, 735-767 (1981).
13. Jacobs, G. K. & Kerrick, D. M. *Geochim. cosmochim. Acta* **45**, 607-614 (1981).
14. Robie, R. A., Hemingway, B. S. & Fisher, J. R. *U.S. geol. Surv. Bull.* **1452**, (1979).
15. Valley, J. W., McLelland, J., Essene, E. J. & Lamb, W. *Nature* **301**, 226-228 (1983).
16. Bohlen, S. R., Peacor, D. R. & Essene, E. J. *Am. Miner.* **65**, 55-62 (1980).
17. Valley, J. W. & O'Neill, J. R. *Geochim. cosmochim. Acta* **45**, 411-419 (1981).
18. Cameron, E. N. & Weis, P. L. *U.S. geol. Surv. Bull.* **1082-E** (1960).
19. Field, D. & Starmer, I. C. *Nature* **298**, 303-304 (1982).
20. Glassley, W. E. *Contr. Miner. Petrol.* **84**, 15-24 (1983).
21. Buddington, A. F. & Lindsley, D. H. *J. Petrol.* **5**, 310-357 (1964).
22. Bohlen, S. R., Essene, E. J. & Hoffman, K. *Bull. geol. Soc. Am.* **91**, 110-113 (1980).
23. Perkins, D., Essene, E. J. & Marcotty, L. A. *Can. J. Earth Sci.* **19**, 1759-1774 (1982).
24. Anderson, A. T. *J. Geol.* **76**, 528-547 (1968).
25. Rollinson, H. R. *Miner. Mag.* **43**, 623-631 (1980).
26. Myers, J. & Eugster, H. P. *Contr. Miner. Petrol.* **82**, 75-90 (1983).
27. Spencer, K. J. & Lindsley, D. H. *Am. Miner.* **66**, 1189-1201 (1981).

Black acidic snow in the remote Scottish Highlands

T. D. Davies, P. W. Abrahams, M. Tranter,
I. Blackwood, P. Brimblecombe & C. E. Vincent

School of Environmental Sciences, University of East Anglia,
Norwich NR4 7TJ, UK

Parts of Scotland experience large acidic depositions^{1,2} which may result in ecological damage³. Some highland areas appear to suffer enhanced acidic input² which frequently exhibits pronounced episodicity⁴. Acidic snowfall leads to acidic flushes during snow-

pack melting^{5,6}. Polluted grey snow layers have been reported in Scandinavia⁷⁻¹⁰, but we present the first documented case remote from major pollution sources of a distinctive black, acid snowfall episode with pH 3.0. The particulate deposit was very large and consisted of ~29% carbon; the remaining fraction included particles which could be identified as coal fly-ash. The heavily polluted air originated to the south-south east and was probably transported at an altitude of ~1,000–1,500 m in association with a stable atmospheric layer whereupon particles were efficiently scavenged by snow. Such transport and deposition mechanisms may produce the greater pollutant deposition sometimes observed in mountain areas². Similar events may not be rare and could make an important contribution to the annual pollutant input.

Shortly before 06.00 on 20 February 1984 a black snowfall (~5 cm) fell in the Cairngorm mountains (data from Aviemore Meteorological Station). Our survey, by road and foot transects in the more accessible zones, indicated that ~200 km² were affected. Individual samples (22) were collected within hours of falling. Because of the difficult conditions, the sample locations were not equally spaced but were biased towards those that appeared to be close to the eastern and western extremities of the fall. The particulate loading was exceptionally heavy (~0.01 g l⁻¹ of melted snow) and our conservative estimate is a deposit of 0.1 tonne km⁻². Detailed analysis has been undertaken on samples collected at 1,080 m, OS National Grid NJ013046. Our survey of snow packs southwards to Edinburgh, a few weeks later, revealed no trace of the event which by then had evolved into a distinctive grey layer ~40 cm thick in the Cairngorms. Local residents claim that black snow has fallen a number of times annually in recent years and appears to favour higher altitudes.

Figure 1a shows the synoptic pattern. For several days preceding the event, a ridge from a strong anticyclone to the east influenced the eastern side of the United Kingdom. Pronounced anticyclonic subsidence inversions were observed (Fig. 2) and at 00.00 on 20 February 1984 the lapse rate in the sub-inversion air was close to the dry adiabatic value, unlike the previous day when it was relatively stable¹². Snow showers, hail showers and cumulonimbus development were also observed in NE Scotland and off the coast at around 06.00 and during the following daylight hours¹². The lower boundary of the black snow layer was a thin stratum of hailstones.

Back trajectories for air over the Cairngorms are shown in Fig. 1b. The air flowed over a broad corridor of industrial and urban activity from the Trent Valley and southern Yorkshire, northwards over Edinburgh and Dundee. Given the uncertainties of trajectory analysis, the large industrial area of Tyneside-Teesside could have contributed, as could smaller zones of activity. European sources could also have contributed to the overall atmospheric pollutant loading, as black atmospheric

Table 1 Composition of black snow and composition range of 12 other snows collected within 2–3 hours of falling during January–April 1984

	pH	SO ₄ ²⁻	Cl ⁻	NO ₃ ⁻	NH ₄ ⁺	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	Charge balance
Black snow	3.0	19.8	14.8	20.9	0.3	0.63	0.52	4.52	0.39	+5.4%
Other snows	3.5–4.7	0.27–4.86	0.15–10.23	0.11–3.52	ND	ND	0.02–0.64	0.07–4.45	ND	ND

Ion concentrations are given in p.p.m.; pH was determined immediately on melting and after filtering through a 0.45-µm Oxoid membrane. Ion determinations were made after storage at 4 °C. Anions were determined on a Dionex Model 12 ion chromatograph (precision ~5%). Cations were measured on a Pye Unicam spectrophotometer (Na⁺, ±0.1; K⁺, Mg²⁺, Ca²⁺, ±0.02) except for NH₄⁺ which was determined by the indophenol blue method on an autoanalyser. There is some evidence that the charge balance is due to organic ions, although pH measurement errors may make a contribution. Charge balance is excess equiv./total equiv. ND, not determined.

particle episodes have been observed in Sweden with winds in the England to Poland sector¹⁹. Pollutant plumes from individual large sources can be tracked for hundreds of kilometres^{20–22} and when they become entrained within or beneath an elevated stable layer there may be little cross-wind dispersion and retention of plume coherence for considerable distances²⁰. The importance of transport above the mixing layer is that the process of dry deposition is inhibited and there is potential for very long range transport²³. Long-range transport of power plant plumes has been found, in the United States, to be favoured at heights of ~800–1,500 m at night when wind speeds were up to ~14 m s⁻¹ at this level; when surface wind speeds were <4 m s⁻¹; when emission was into the lower atmosphere with a dry adiabatic lapse rate and the plume became confined in a more (and becoming more) stable atmosphere²⁴. These atmospheric conditions show strong similarities with those prevailing in this case¹². Such stable plumes have been reported to be dry-deposited on high terrain at night²⁵.

The elevation of the mountains, the geography to the south, black snow distribution and the requirement for little-dispersed, heavily-polluted air to be available for scavenging suggest that advection of the greater proportion of the pollution was within or just below the subsidence inversion layer. This implies that the pollution would have to rise to about 1,000 m and implicates large emitters producing buoyant plumes, although low-level emissions may have made a contribution (synoptic observations of smoke and haze indicated a northward transport of pollution in the lower layers; Fig. 2). Power station plumes, for example, can rise to these heights under the light wind conditions²⁶ which prevailed during the previous night¹¹. Figure 2 shows the heavily polluted air being made available for snow scavenging over the Cairngorms. It was clear that snow scavenging, and not dry deposition, was the removal process as the particles were

Fig. 1 *a*, Synoptic pattern at 06.00 on 20 February 1984¹¹ with locations referred to in Fig. 2 (L. Lerwick; S. Shanwell; H. Hemsby). *b*, Back-trajectories from the Cairngorms at 00.00 (●) and 06.00 (—) on 20 February 1984, at 6-hour intervals. The geostrophic approximation for the surface pressure distribution was used. The plan of the 00.00 trajectory agrees well with the observed 850 mbar winds¹² although use of these would increase the length of the 6-hour trajectory by 40%. There are considerable uncertainties and debate over trajectory analysis^{4,13}. The 850 mbar wind has been used¹⁴ and the use of the surface pressure field is considered to be robustly appropriate^{15–18}. In this case the cross-isobar component of flow was restricted to a very shallow layer of air near the ground¹². Both trajectories are sufficiently similar for us to be confident (in the absence of an exact time for the event) of the likely pollutant source regions.

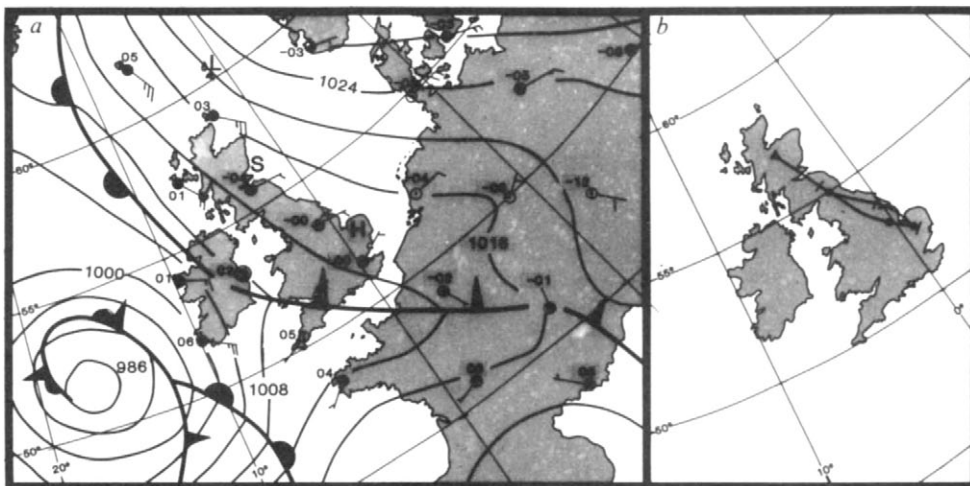
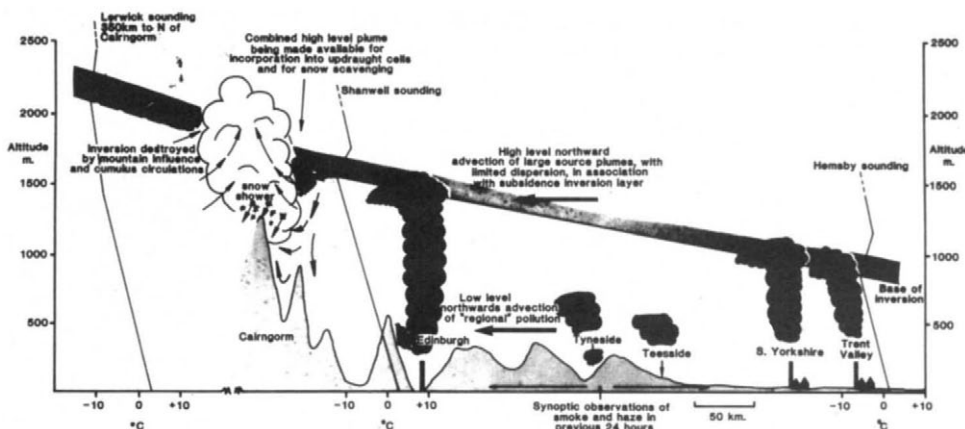


Fig. 2 Schematic cross-section along back-trajectory to the Trent Valley. Shown are dry-bulb temperature soundings at 00.00 on 20 February 1984 (see Fig. 1). Tyneside and Teeside are industrial areas to the east of the probable trajectory, but may have contributed to the 'regional' pollution plume. Advection in association with the inversion layer made available little-dispersed, heavily-polluted air for scavenging. Pollution in the less stable, sub-inversion layer was more efficiently dispersed and diluted and further depleted by dry deposition, especially in hilly terrain. The strength and position of the inversion fluctuated temporally and spatially¹²; consequently the northward advection of the inversion-layer pollution should not be regarded as increasing in height in the simple fashion portrayed. No showers occurred along the trajectory south of the Scottish mountains¹¹ and so high pollutant concentrations in the mountain air were available for scavenging via down-mixing induced by topographic disturbance, and compensating downdraughts associated with the cumulus circulations. These cumulus circulations may have drawn in volumes of the heavily polluted air, effectively 'concentrating' the pollution to produce a pronounced localized deposition²⁶.



incorporated throughout the whole depth of the fall. Snow is a much more efficient scavenger of aerosols than rainfall^{27,28} and scavenging efficiency increasing with decreasing ice-crystal diameter²⁹. This episode consisted of fine-grained snow.

The black snow had a pH of 3.0 at the main sampling site. The volume-weighted mean annual precipitation in this area of Scotland is around 4.5 (ref. 2), although the pH is lower for our recent snow samples from greater elevations (Table 1). Of particular interest are the relatively large excess of Cl^- and the NO_3^- concentrations in the black snow. HCl is very soluble and so will be relatively easily removed once precipitation occurs. NO_x oxidizes before SO_2 (ref. 30) and snow is relatively efficient at removing NO_3^- (refs 28, 29). The molar ratios of NO_3^- : non-sea salt SO_4^{2-} : non-sea salt Cl^- (1.71:1.00:1.04) may therefore reflect pollution plume ageing, different solubilities and differential scavenging, or different sources.

Fog or cloud droplet deposition may increase acidic input at this altitude³¹, but our observations appear to show a systematic increase of pH with height (Table 2), suggesting that this was not the most important mechanism in this case. Snowflakes at lower elevations had a longer fall-path through polluted air mixed down into the mountains from high level (Fig. 2) or, if the cloud enveloped the mountains, the mechanism may have been early scavenging of ascending polluted air drawn into the cloud.

Figure 3 shows many particles which are typical of pulverized fuel (PF) or coal fly ash^{32,33}. The carbonaceous component, which gave the snowfall its distinctive appearance, comprises much of the particulate volume but it is highly porous (with a specific area of up to 80 times that of non-carbonaceous oil-fired fly ash³⁴). The carbon content of the particulate is $29 \pm 1.5\%$ by weight (Hewlett-Packard 185 B CHN analyser). The carbon content of chimney solids from modern UK oil-fired power stations is $<5\%$ by weight³⁴, but values in excess of 20% have been reported in the United States under some operating conditions^{32,35}. Other coal burners produce larger proportions of carbon residue and would have contributed to the black snow carbonaceous component. For example, the carbon content of domestic coal ash is very high³⁴. The carbonaceous residue particles in PF fly ash are often concentrated in the fraction greater than $45 \mu\text{m}$ in diameter³⁴, and as they are highly porous, they will become relatively enriched at the expense of other, heavier fly ash components which settle out much more rapidly in the long-range transport of the plume. In addition, the large surface area of the carbonaceous particles will lead to further enrichment in the snow scavenging process. The carbonaceous residue may be of particular interest to the acidic deposition issue as it has a large capacity for absorbing acids in the plume. The pH of a soot/water slurry decreases with time whereas that of a non-carbonaceous PF fly ash/water slurry increases³⁴.

Besides glassy spheres and the very porous carbon background, other less easily identifiable particles are seen in Fig. 3 and these may be from a variety of sources. Elemental proportions (Table 3) are similar to those of PF fly ash^{35,36} (and perhaps to other industrial ashes though data are not available) although Ca proportions are relatively low (the Ca fraction is soluble, and there are high concentrations in the meltwater) and Fe is higher than 'typical' but well below the 'common maximum'³⁷. The Pb, Mn and Zn concentrations (respectively $1,530 \pm 11$, $4,960 \pm 94$, 9830 ± 108 p.p.m. after HNO_3 and HClO_4 digestion and atomic absorption spectrophotometry) are higher than those typically observed in PF chimney solids³⁸ (especially Zn) but trace elements, in particular Zn³⁹, are known to be enriched on small particles⁴⁰⁻⁴³ which have much longer atmospheric residence times. The log probability size distribution (Malvern 3600 E-type particle size VA3 analysis after ultrasonic dispersion) for the fraction $<4 \mu\text{m}$ diameter ($\sim 70\%$ of the total number) was similar to distributions determined for PF chimney solids⁴⁰, but shifted towards smaller diameters. Our distribution was deficient in $>4 \mu\text{m}$ particles, presumably as a result of preferential removal in the ageing atmospheric plume (this is not inconsistent with carbonaceous component enrichment in the plume when the weight of carbon as a proportion of different size classes of fly ash from PF power stations is considered³⁴).

In the region of the back trajectory there is a total of $\sim 20,000$ MW of PF power station capacity, the largest plant in

Table 2 Snowfall pH variation with altitude on a hill 18 km to the west of Cairngorm

Altitude (m)	280	295	320	410	500	570
pH	2.95	3.00	3.10	3.15	3.30	3.30

It is possible that this variation reflects an 'edge' effect as this was the most westerly sampling point and altitude increased towards the west.

Table 3 Range of percentage contribution of measured elements to the total weight of all 14 constituent elements in black snow samples

Na	2.2-5.4	Ca	0.6-1.1
Mg	4.0-7.5	Ti	0.5-1.3
Al	18.6-24.0	V	0.0-0.1
Si	30.3-44.3	Cr	tr-0.1
S	1.3-1.8	Fe	18.9-21.4
Cl	0.3-0.4	Ni	0.1-0.15
K	1.9-5.1	Cu	0.05

Values obtained by EDAX bulk analysis⁴⁰ over areas of 20 mm^2 of membrane filters (three analyses). The accuracy of determination for each element is around $\pm 5\%$. tr = trace.

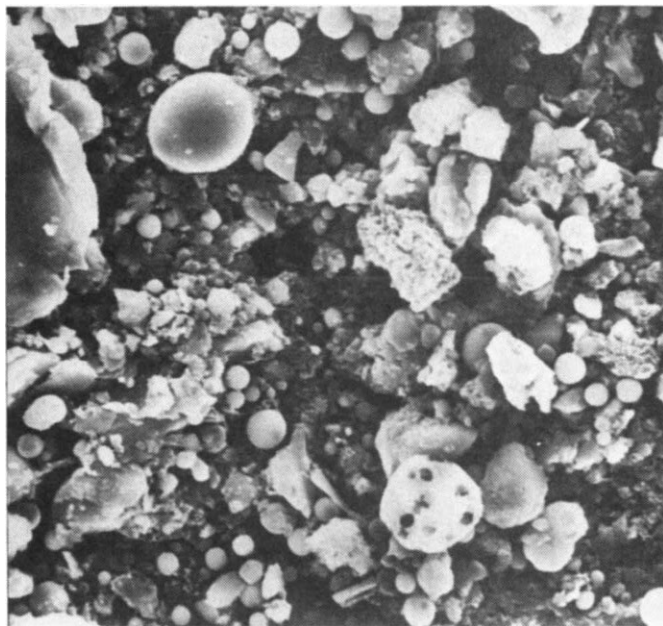


Fig. 3 Scanning electron microscope photograph of particulate material which shows many glassy spheres generally $<5\ \mu\text{m}$ diameter. The spongy, lattice-like background is carbon residue. The large sphere to top left is $\sim 8\ \mu\text{m}$ diameter. The cenosphere (porous particle near bottom) may be tarry, coal exudate but is a typical product of oil-firing (they are relatively infrequent in our samples, ~ 1 in 200).

the Trent Valley and southern Yorkshire. Other large buoyant sources include chemical plant near Teesside and an oil refinery (with its own 100-MW oil-fired power station) to the west of Edinburgh (although our trajectory analysis favours a route to the east of the city). Areal sources in cities may have also been sufficiently buoyant to become entrained in the high-level advection. No information is available on the nature of emissions from the refinery although large quantities of SO_2 , NO_x and organics are released from plant in the United States⁴⁴. The refinery particulate discharge is likely to be small⁴⁴ compared to the 20 tonnes per day suggested for a modern UK 1,800-MW PF power station (generating at full load and assuming an ash capture efficiency of 99.3%)³⁴. It is expected that oil ash would contain much higher V and Ni concentrations than our samples (166 ± 50 , 176 ± 10 p.p.m. respectively) and more cenospheres⁴⁵.

Whatever the major source(s) of this large pollutant deposit it is likely that disparate activities have contributed. The exposed faces of the evolved grey snow layer produced a distinctive odour which pervaded the mountainside under calm conditions. It was difficult to characterize the smell and to identify the agents, but some observers have described it as an acetone/phenol-type odour. Petrochemical sources may have contributed. The heavy particulate deposit was the most remarkable feature of this episode. For example, it would require the entire emission for more than two hours from $\sim 20,000$ MW of efficiently-operating PF power stations³⁴.

Such severe acidic episodes could enhance the spring snow pack 'acid flush', both through heavy loading and the 'banding' of acidity in the snowpack (A. R. W. Marsh and A. H. Webb, unpublished CEGB report, 1978). It is of interest to note that the weather pattern was similar to that which produces acidic episodes in Scandinavia⁴⁶, from widespread sources. There is some evidence that the frequency distribution of this type is changing on time scales from years to decades (T. D. Davies and P. M. Kelly, unpublished work) and there may be implications for dry as well as wet deposition in polluted air volumes retaining coherence over long distances and becoming available for transfer to the ground in mountain areas²⁵.

We are grateful for the financial support of the NERC and EEC. We thank Dr A. G. Thomas, Institute of Hydrology, for

assistance with NH_4^+ determinations, and the Central Electricity Generating Board for helpful comment and discussion.

Received 18 June; accepted 12 September 1984.

1. Fowler, D. *et al.* *Nature* **297**, 383–386 (1982).
2. UK Review Group on Acid Rain *Acid Deposition in the UK* (Warren Spring Laboratory, Stevenage, 1983).
3. Leivestad, H. & Muniz, I. P. *Nature* **259**, 391–392 (1976).
4. Smith, F. B. & Hunt, R. D. *Atmos. Environ.* **12**, 461–477 (1978).
5. Davies, T. D., Vincent, C. E. & Brimblecombe, P. *Nature* **300**, 161–163 (1982).
6. Johannessen, M. & Henriksen, A. *Wat. Resources Res.* **14**, 615–619 (1978).
7. Henriksen, A. *Vatten*, 409–412 (1972).
8. Elgmork, K., Hagen, A. & Langeland, A. *Environ. Pollut.* **4**, 41–52 (1973).
9. Hagen, A. & Langeland, A. *Environ. Pollut.* **5**, 45–57 (1973).
10. Wright, R. F. & Dovland, H. *Atmos. Environ.* **12**, 1755–1768 (1968).
11. Daily Weather Summary, London Weather Centre.
12. European Weather Bulletin, Offenbach am Main, W. Germany.
13. Sykes, R. I. & Hatton, L. *Atmos. Environ.* **10**, 925–934 (1976).
14. *OECD Programme on Long Range Transport of Pollutants* (OECD, Paris, 1977).
15. Munn, R. E. & Rodhe, H. E. *Tellus* **23**, 1–13 (1971).
16. Kurtz, J. & Schneider, W. A. *Atmos. Environ.* **15**, 1111–1116 (1981).
17. Forland, E. J. *Tellus* **25**, 291–298 (1973).
18. Fowler, D. & Cape, N. *Atmos. Environ.* **18**, 1859–1866 (1984).
19. Brossett, C. *Ambio* **5**, 157–163 (1976).
20. Oke, T. R. *Boundary Layer Climates* (Methuen, London, 1978).
21. Cocks, A. T., Kallend, A. S. & Marsh, A. R. W. *Nature* **305**, 122–123 (1983).
22. Brimblecombe, P. & Davies, T. D. *Weather* **33**, 19–27 (1978).
23. Fisher, B. E. A. in *Sulfur in the Environment* (ed. Nriagu, J. O.) 243–296 (Wiley, New York, 1978).
24. Smith, T. B., Blumenthal, D. L., Anderson, J. A. & Vanderpol, A. H. *Atmos. Environ.* **12**, 605–611 (1978).
25. Long, J. H. & Williams, D. R. *Aerial Techniques for Environmental Monitoring, Tropical Symp. Nuclear Soc.*, Las Vegas, Nevada (7–11 March, 1977).
26. Pasquill, F. & Smith, F. B. *Atmospheric Diffusion* (Horwood, Chichester, 1983).
27. Knutson, E. O., Sood, S. K. & Stockham, J. D. *Atmos. Environ.* **10**, 395–402 (1976).
28. Raynor, G. S. & Hayes, J. V. *Atmos. Environ.* **16**, 1647–1656 (1982).
29. Sood, S. K. & Jackson, M. R. in *Precipitation Scavenging* (1970) 121–136 (US Atomic Energy Commission, National Technical Information Service, Virginia, 1970).
30. Rodhe, H., Crutzen, P. & Vanderpol, A. *Tellus* **33**, 132–141 (1981).
31. Dollard, G. J. & Unsworth, M. H. in *Precipitation Scavenging, Dry Deposition and Resuspension* (eds Pruppacher, H. R. *et al.*) 161–170 (Elsevier, New York 1983).
32. McCrone, W. C. & Delly, J. G. *The Particle Atlas* Vol. 3 (Ann Arbor, Michigan, 1973).
33. Stern, A. C. *Air Pollution* Vol. 2 (Academic, New York, 1968).
34. Raask, E. *Power Ind. Res.* **1**, 233–251 (1981).
35. Stern, A. C. *Air Pollution* Vol. 3 (Academic, New York, 1968).
36. Raask, E. *J. Inst. Energy* **53**, 70–75 (1980).
37. Gilham, E. W. F. *Trace Elements in Soils and Ashes—A Comparative Review* (2), CEGB MID/SSD/81/0029/R (1981).
38. Green, D. E. *CEGB Rep. SSD/NE/R380* (1979).
39. Klein, D. H. *et al.* *Environ. Sci. Technol.* **9**, 973–979 (1975).
40. Raask, E. & Goetz, L. *J. Inst. Energy* **54**, 163–173 (1981).
41. Ondov, J. M., Raigani, R. C. & Biernann, A. H. *Atmos. Environ.* **12**, 1175–1185 (1978).
42. Raigani, R. C. & Ondov, J. M. in *International Conference on Environmental Sensing and Assessment* Vol. 1 Section 17-2, Las Vegas, Nevada (1975).
43. Davison, R. L., Natusch, D. F. S., Wallace, J. R. & Evans, C. A. *Env. Sci. Technol.* **8**, 1107–1113 (1973).
44. Sexton, K. & Westburg, H. *J. Air. Poll. Cont. Ass.* **29**, 1149–1152 (1979).
45. Hart, A. B. & Lawn, C. J. *CEGB Res.* **5**, 4–17 (1977).
46. Smith, F. B. *Acid Precipitation Conf.* (Lindau, FGR, 1983).

Cell size control of development in *Saccharomyces cerevisiae*

Geoffrey R. Calvert & Ian W. Dawes

Department of Microbiology, University of Edinburgh, Edinburgh EH9 3JG, UK

Diploid cells of the yeast *Saccharomyces cerevisiae* in the G_1 phase of the cell cycle are faced with the alternatives of either continuing vegetative cell division or undergoing the developmental processes of meiosis and subsequent ascospore formation, or adapting to starvation conditions if these apply. The course taken is influenced by the nutritional status of the culture medium¹, the presence of both *MATa* and *MAT α* mating-type alleles², and the need for cells to be in the G_1 phase of the cell cycle³. For those cells that continue cell division, size controls operate in both the budding yeast *S. cerevisiae*^{4,5} and the fission yeast *Schizosaccharomyces pombe*^{6,7}. In *S. cerevisiae* the 'start' event initiating the cell cycle is controlled in some way related to cell size because cells below a critical size fail to initiate cell division⁸. The ability of cells to undergo the developmental process of sporulation is related to cell age, in that cells gain this ability just before the emergence of the first bud⁹ and the process of sporulation after initiation is altered in small cells¹⁰. Here we report that the initiation of sporulation is subject to a size control related to absolute cell volume, which is distinct from control by cell age and also independent of the control operating on the initiation of cell division.

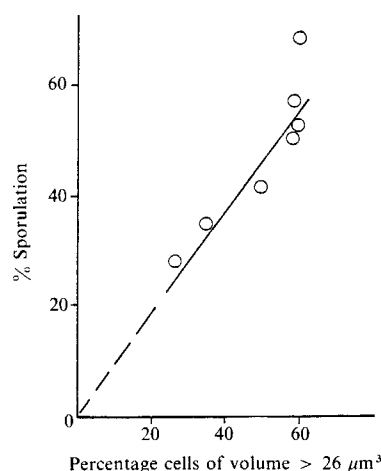


Fig. 1 Relationship between the ability of a yeast population to sporulate and the percentage of cells of volume $>26\ \mu\text{m}^3$ before the shift to sporulation medium. Cells of strain S41 grown aerobically in continuous culture at dilution rates between $0.02\ \text{h}^{-1}$ and $0.09\ \text{h}^{-1}$ were examined microscopically and individual cell volumes measured. Cells were collected from 30 ml samples, washed and resuspended in sporulation medium ($20\ \text{g l}^{-1}$ potassium acetate). After 48 h incubation at 30°C with shaking, the sporulation frequency of the samples was measured. Further samples were taken after 72 h to check that sporulation was complete. The chemostat medium contained (per l): 10 g galactose, 5 g ammonium sulphate, 1.5 g Difco yeast nitrogen base, 1 g Oxoid yeast extract, 0.1 g arginine, 5.5 g potassium dihydrogen phosphate, 10 g dipotassium hydrogen phosphate and amino acids at 0.1 g, pH 7.2.

Initially, it was noted in strains carrying the *whi1* or *whi2* size control mutation^{11,12} alone or in combination with the *spd1* mutation which derepresses sporulation¹³, that smaller cells in the population did not sporulate whereas the few larger cells present did¹⁴. These preliminary experiments suggested that *whi1* homothallic diploids were unable to sporulate, but did not indicate whether this was a direct effect of the *whi1* mutation, or simply an effect of the smaller cell size. This was tested by growing cells in a presporulation medium containing hydroxyurea, an inhibitor of DNA synthesis which arrests nuclear division without preventing mass growth, and therefore causes cells to become enlarged¹⁵. The results are shown in Table 1: *whi1* cells which were normally unable to sporulate could form large numbers of asci when made larger by preincubation with hydroxyurea. This indicates that the *whi1* mutation was not directly preventing sporulation, but was affecting the process through its property of reducing cell size. These results, while indicating that there may be some size control over initiation of sporulation, are subject to other interpretations as mutants were used.

By growing wild-type yeast at slow growth rates in continuous culture, very small cells can be obtained, and cultures with widely varying average cell sizes can be produced by changing the dilution rate¹⁶. We have studied the dependence of sporulation on cell size by growing populations in chemostat cultures over a wide range of dilution rates. Cultures were grown in media in which sporulation is not initiated but in which mitochondrial functions that are essential for spore formation are derepressed. At each steady state the distribution of cell sizes in the cultures was estimated microscopically and samples of the culture were transferred to sporulation medium to assess the extent to which that population of cells was competent to sporulate. From the results, we estimated the average cell size at which half the maximum sporulation frequency was obtained (that is, 40%); this gave an estimate of the critical volume for initiation of sporulation of $26\ \mu\text{m}^3$. Figure 1 illustrates the linear relationship between the extent to which a culture sporulated on resuspension in sporulation medium and the proportion of cells of volume $>26\ \mu\text{m}^3$ before the shift to sporulation medium.

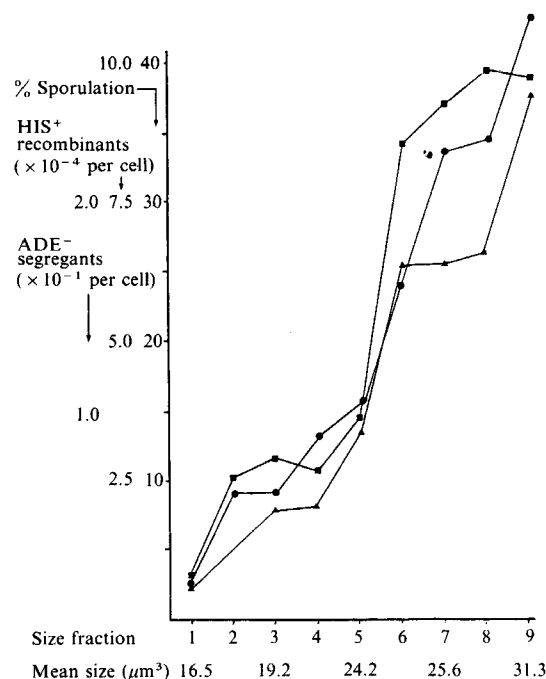


Fig. 2 Cells of strain 60x61,

MATα HO HMLα HMRA ade2-40 his4-239 can1 CYH2 ARG4
MATα HO HMLα HMRA ade2-119 his4-166 CAN1 cyh2 arg4-1

were grown in the chemostat medium described in Fig. 1 legend. This strain requires histidine for growth but not adenine. On meiotic recombination, however, it yields histidine prototrophs and on chromosome segregation it yields red adenine auxotrophs. Cells were collected from 10 ml of chemostat culture, washed in sporulation medium, and resuspended in the least possible volume. The suspension was loaded onto a gradient of Urografin (density $1.04\text{--}1.08\ \text{g cm}^{-3}$, made up in 2% potassium acetate) in 16.5-ml tubes, and centrifuged for 13 min at 60 g. After centrifugation, the part of the gradient containing the cells was split into nine fractions and the cells from each fraction were collected and resuspended in 3 ml of sporulation medium. Samples of each fraction were diluted appropriately and plated onto minimal medium ($1.5\ \text{g l}^{-1}$ Difco yeast nitrogen base, $5\ \text{g l}^{-1}$ $(\text{NH}_4)_2\text{SO}_4$, $20\ \text{g l}^{-1}$ glucose, $20\ \text{g l}^{-1}$ agar) and YEPD ($10\ \text{g l}^{-1}$ yeast extract, $20\ \text{g l}^{-1}$ peptone, $20\ \text{g l}^{-1}$ glucose, $20\ \text{g l}^{-1}$ agar) immediately after resuspension and after 48 h incubation at 30°C . The figure shows the percentage sporulation (●), appearance of histidine prototrophs (▲), and appearance of adenine red auxotrophs (■) in each size fraction.

These data could be interpreted in terms other than the existence of a critical cell size control over initiation of sporulation as: first, they only refer to the production of mature ascospores and therefore cells may be initiating, but not completing sporulation; and second, they only deal with average cell sizes of populations, and not with individual cells. The first of these possibilities was excluded by examining the frequency of intragenic recombination at the *HIS4* locus in populations of sporulating cells of different sizes obtained from a single parent culture by density gradient fractionation. The results (Fig. 2) show that small cells were not undergoing meiotic levels of intragenic recombination at the *HIS4* locus or chromosome segregation. Thus, small cells do not proceed as far as the early sporulation event of intragenic recombination.

The sporulation of individual cells was examined by growing cells of strain 60x61, heterozygous at the *ADE2* locus, in a medium giving a range of cell sizes, resuspending them in sporulation medium, and after 20 h of incubation, dissecting out 250 cells of $<30\ \mu\text{m}^3$ volume, and 256 cells of $>40\ \mu\text{m}^3$ volume. On subsequent incubation, 110 of the latter were found to form red or sectorial colonies, indicating segregation at the *ADE2* locus, while none of the former showed segregation. We were able to measure cell sizes after 20 h in sporulation medium

Table 1 Hydroxyurea restores the ability of *whi1* mutants to sporulate

	Mean cell volume (μm^3)	% Ascus formation
Wild type	78.2	65.4
Wild type (HU-treated)	ND*	62.3
<i>whi1</i>	22.1	1.6
<i>whi1</i> (HU-treated)	50.4	29.0

Cells of strains S41: *MATa HO HMLa HMRa arg-1 cyh1*
MATa HO HMLa HMRa arg-1 cyh1
 and 65: *MATa HO HMLa HMRa whi1 met5*
MATa HO HDLa HMRa whi1 met5

were inoculated into YEPAG medium (10 g l⁻¹ yeast extract, 20 g l⁻¹ peptone, 20 g l⁻¹ potassium acetate, 1 g l⁻¹ galactose), and after 2 h incubation, hydroxyurea HU; 50 mM) was added to half the samples and incubation was continued. When the absorbance at 600 nm reached 1.0, the cell size distribution of each culture was measured microscopically, and the cells were transferred to sporulation medium. After 48 h in sporulation medium, the frequency of asci in each culture was measured. ND, not determined.

* Large elongate cells of irregular shape.

as cell size in this strain did not significantly increase during sporulation.

The sporulation of individual cells was also examined by photographing 350 cells before and after sporulation on solid medium. No cells smaller than 25 μm^3 sporulated, and the probability of sporulation increased with increasing size above 25 μm^3 , up to a maximum frequency approaching 100% in cells larger than 130 μm^3 .

The cell size control over sporulation was distinguished from the cell age control by examining the sizes at which cells were able to initiate budding. Small cell fractions (obtained by methods described in Fig. 2 legend) were resuspended in rich growth medium and the sizes of budded cells were measured. The minimum size of budded cells was ~17 μm^3 , and about 8% of budded cells were <26 μm^3 in volume. This indicates that some cells of <26 μm^3 volume were capable of budding, therefore the inability of cells smaller than 26 μm^3 to sporulate was not simply due to their being unbudded daughter cells.

We have used a variety of different methods which give consistent results: yeast cells with a volume less than ~26 μm^3 are unable to sporulate, and this effect is unrelated to size control over budding and cell age control over sporulation. These results have interesting implications for other developmental systems. Size (or critical cell mass) control of cell division seems to operate in higher eukaryotes¹⁷ and similar effects may therefore influence the course of higher cell development.

This work was supported by grants from the SERC. We thank P. Sudbery, B. L. A. Carter and J. R. Dickinson for helpful discussion.

Received 22 June; accepted 21 August 1984.

- Esposito, R. E. & Klapholz, S. in *The Molecular Biology of the Yeast Saccharomyces cerevisiae: Life Cycle and Inheritance* (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 211-297 (Cold Spring Harbor Laboratory, New York, 1981).
- Roman, H. & Sands, S. M. *Proc. natn. Acad. Sci. U.S.A.* **39**, 171-179 (1953).
- Haber, J. E. & Halvorson, H. O. *J. Bact.* **109**, 1022-1033 (1972).
- Johnston, G. C., Erhardt, C. W., Lorincz, A. & Carter, B. L. A. *J. Bact.* **137**, 1-5 (1979).
- Wheals, A. E. *Molec. cell. Biol.* **2**, 361-368 (1982).
- Fantes, P. & Nurse, P. *Expl Cell Res.* **107**, 377-386 (1977).
- Thuriaux, P., Nurse, P. & Carter, B. L. A. *Molec. gen. Genet.* **161**, 215-220 (1978).
- Johnston, G. C., Pringle, J. R. & Hartwell, L. H. *Expl Cell Res.* **105**, 79-98 (1977).
- Sando, N., Maeda, M., Endo, T., Oka, R. & Hayashibe, M. *J. gen. appl. Microbiol.* **19**, 359-373 (1973).
- Ueki, K., Yoshimi, S. & Tanji, T. *J. gen. appl. Microbiol.* **27**, 1-9 (1981).
- Carter, B. L. A. & Sudbery, P. E. *Genetics* **96**, 561-566 (1980).
- Sudbery, P. E., Goodey, A. R. & Carter, B. L. A. *Nature* **288**, 401-404 (1980).
- Dawes, I. W. *Nature* **255**, 707-708 (1975).
- Calvert, G. R. thesis, Univ. Edinburgh (1983).
- Slater, M. L. *J. Bact.* **113**, 263-270 (1973).
- Lorincz, A. & Carter, B. L. A. *J. gen. Microbiol.* **113**, 287-295 (1979).
- Killander, D. & Zetterberg, A. *Expl Cell Res.* **40**, 12-20 (1965).

New evidence that growth in 3T3 cell cultures is a diffusion-limited process

Graham A. Dunn & Grenham W. Ireland

MRC Cell Biophysics Unit, 26-29 Drury Lane,
London WC2B 5RL, UK

When a confluent culture of 3T3 cells is wounded, new growth occurs at the wound margins. This indicates that the suppression of growth in the intact confluent sheet is under local control, a phenomenon known as 'topoinhibition'¹, and it has been suggested that intercellular contact is responsible. An alternative explanation for topoinhibition is possible, however, namely that a soluble factor necessary for growth is locally depleted from the medium by cells so that each cell in a confluent sheet normally receives an insufficient supply and its growth is inhibited²⁻⁵. Here we show that the pattern of release from topoinhibition in a wounded culture can be distorted simply by inducing a gentle laminar flow of medium across the wound. Growth remains suppressed at the upstream margin of the wound despite the reduced level of intercellular contact at wound edges. We conclude that the signal for topoinhibition is carried by the flow as would be predicted if it were the local depletion of growth factor.

Hydrodynamic theory predicts the existence of a thin layer of non-turbulent medium next to the cells in which diffusion is the main mechanism of transport. Any growth factor consumed by the cells would thus become locally depleted in this boundary layer. This explanation of topoinhibition is reinforced by the discovery that forced turbulent stirring of the culture medium using either a miniature pump³ or vigorous shaking of the cultures⁴ increases the growth rate of 3T3 cells. Stirring reduces the thickness of the boundary layer and could, therefore, increase the diffusive flux of growth factor to the cells³⁻⁵; however, this interpretation has been criticized on the grounds that stirring could disrupt intercellular contacts and so the alternative hypothesis of growth control by contact is not eliminated⁶. No effects of fluid shear stress on intercellular contacts have been found although shear stress does affect the migration⁷ and cytoskeleton⁸ but not the growth⁷ of endothelial cells in culture. A further criticism is that increasing the viscosity of the medium does not have the effect of reducing the saturation density of 3T3 cultures that might be expected if growth depended on diffusive flux⁶. Thus, we have developed a direct test of whether the release from topoinhibition at wound margins depends on a diffusible factor.

3T3 cells were grown in Petriperm culture dishes in order to maintain steady conditions of pH and gaseous exchange. Confluent monolayers were wounded by scraping away a narrow strip of cells along a diameter of each dish and the medium was replaced with fresh medium containing ³H-thymidine and varying concentrations of serum. In some cultures, the lids were replaced with the apparatus shown in Fig. 1 and the cultures were all incubated for a further 16 h then prepared for

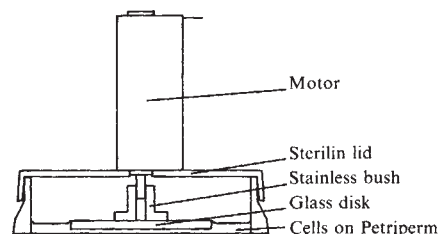


Fig. 1 Diagrammatic section of the apparatus. The motor (Portescap Ltd, type M915L, ratio 1:2,430), which is powered by a constant voltage source, is attached to a Sterilin lid which provides a better location on the Petriperm dish than a Petriperm lid. The 32-mm diameter glass disk is glued to the stainless steel bush and adjusted along the motor spindle to give a separation of 0.8 mm from the Petriperm membrane.

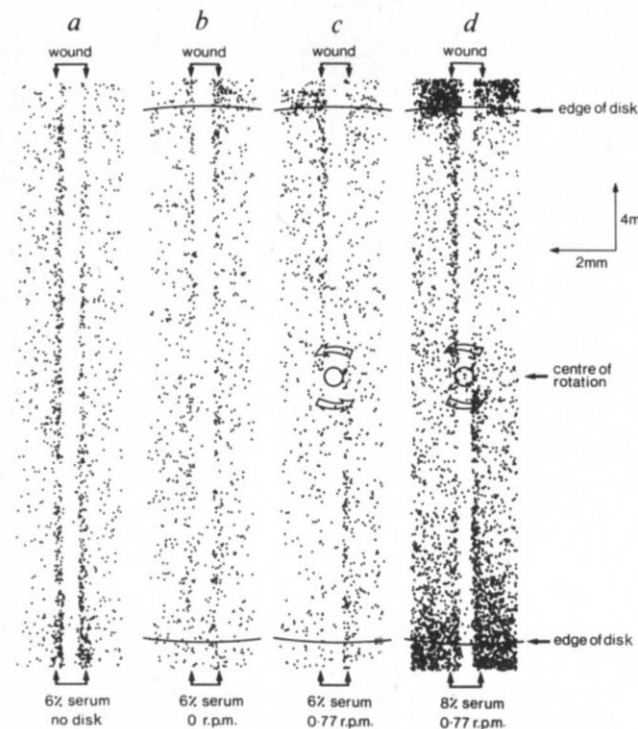


Fig. 2 Effect of different flow regimes and serum concentrations on the distribution of labelled nuclei at wound edges. Swiss 3T3K cells (obtained from Dr C. O'Neill) were grown in bicarbonate-buffered Dulbecco's modified Eagle's medium (Flow) containing 10% fetal calf serum (Seralab) in a 37°C humidified incubator gassed with 15% CO₂ in air. 3×10^5 cells were seeded in 5 ml of medium in 5 cm-diameter Petriperm dishes (Heraeus). After 60 h the medium in the cultures was replaced and on day 5 the cultures were used for the experiment, by which time they had been confluent for 2 days. Precise parallel-sided diametral wounds, 0.9 mm wide, were made with a stainless steel tool guided by a jig. The medium was replaced with 3 ml of fresh medium containing $1 \mu\text{Ci ml}^{-1}$ ³H-thymidine (Amersham International; specific activity 112 Ci mmol⁻¹). In *b*, *c* and *d* the original Petriperm lids were replaced with Sterilin lids bearing motors and the cultures returned to the incubator for 16 h. The cultures were fixed in ice-cold 5% trichloroacetic acid, rinsed in 95% ethanol and air-dried. Autoradiography was performed by attaching stripping film to the cultures, exposing them for 3 days, then developing and fixing. The grain density over labelled nuclei was so high that individual grains could not be distinguished and labelled nuclei could be identified using digital image analysis. A Zeiss RA microscope equipped with luminar objectives and a video camera was used to input a frame showing part of the wound to a digital image analysis system. The background was subtracted and windows for size, shape and grey level were set to identify labelled nuclei. The scale used for plotting nuclear positions is foreshortened along the axis parallel with the wound, as indicated by the two scale bars.

autoradiography. Information about the different flow patterns was obtained in separate experiments using suspended latex spheres of 2 µm diameter.

In the cultures with unmodified lids, the medium shows the slow convective stirring that is thought to be commonly present in undisturbed cultures². Enhanced incorporation of ³H-thymidine at the wound edges is very marked in 4% or 6% serum (Fig. 2*a*). Most of the edge enhancement is contributed by cells that have migrated into the original wound area. With 8% or 10% serum, the background labelling of cells over the whole cell sheet becomes so high that the edge effect is not discernible.

In the cultures with modified lids but with the glass disk kept stationary throughout the experiment, convection currents are suppressed beneath the disk and the movement of latex spheres appears to be confined to brownian motion. At 6% serum concentration (Fig. 2*b*), the enhanced labelling of cells at the

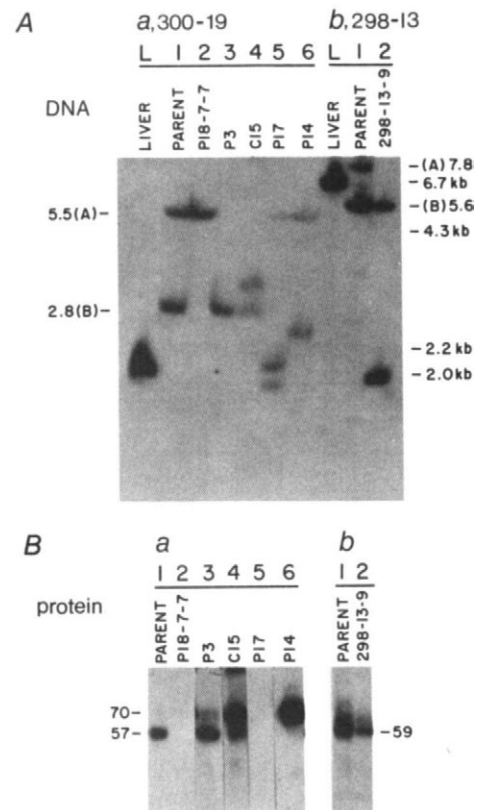


Fig. 3 *a*, Specimen fixed immediately after wounding and stained with Giemsa to show the original parallel-sided wound in a confluent sheet. *b*, Part of the same specimen as Fig. 2*d* but stained with Giemsa. Arrows show direction of flow; most labelled nuclei can be seen on the downstream edge. The extent of migration over the 16 h experimental period can be seen by comparison with *a*. Scale bar, 100 µm.

wound edges is not so marked as in the unmodified cultures. At higher serum concentrations, background labelling beneath the disk is lower than in the unmodified cultures. The extent of cell migration into the wound area does not appear to be affected by the presence of the disk.

If the disk is rotated at 0.77 r.p.m. throughout the experiment, the latex spheres follow concentric circular streamlines, the flow beneath the disk appearing to be entirely tangential with no radial or vertical component. Extensive flow studies with enclosed rotating disks⁹ confirm that this laminar flow is to be expected with our conditions, the velocity varying linearly across the gap from zero at the Petriperm membrane to ωr at the glass disk (where ω is the angular velocity and r is the radial distance from the axis of rotation). With 6% or 8% serum concentration, we see a clear difference in the number of labelled cells at the upstream and downstream edges of the wound. At 6% serum concentration (Fig. 2*c*), the edge effect is suppressed at the upstream edge and, in most cases, it ceases to be noticeable. With 8% serum (Figs 2*d*, 3*b*), the upstream edge effect is noticeable but is still markedly smaller than the downstream edge effect. This asymmetry in labelling of the two wound edges cannot be accounted for by differences in the extent of cell migration into the wound area: the difference in labelling index (proportion of labelled cells) between sampled areas from the upstream and downstream edges within the confines of the original wound area is highly significant (8% serum, upstream: 76/685; downstream: 215/913, $\chi^2 = 41$, d.f. = 1, $P < 0.001$).

The level of serum needed to stimulate new growth beneath a stationary or rotating glass disk is higher than in the absence of the disk. It is unlikely that the disk affects gaseous exchange to the cells since the cells are grown on gas-permeable membrane. The fact that the glass disk, even when rotating, suppresses the vertical component of flow in the medium, suggests

that this reduced growth rate is compatible with the explanation of increased growth caused by shaking cultures⁴. The alternative contact hypothesis does not readily explain these effects.

In the laminar flow experiments, the difference in growth between the two wound edges shows that the normal growth response to local cell density can become distorted by flow. This difference cannot be attributed to a difference in shear stress as this is the same for both wound edges at any particular distance from the centre of rotation. In any case, the maximum shear stress in our system is only 0.012 dyne cm⁻², more than 100 times lower than the stresses reported to affect cells^{7,8}. Furthermore, the extent of cell spreading, migration and loss of contact is similar at the two edges (Fig. 3) and so it is unlikely that the difference in growth is a secondary consequence of flow affecting

these properties. The only difference in local conditions between the two wound edges is that at the upstream edge the medium flows off the cell sheet, whereas at the downstream edge it flows off the cell-free wound area. We conclude that the signal for determining cell growth is carried by the flow so that the growth response at any location is influenced by the cell density extending some way upstream. The most likely candidate for a signal of cell density that can be transferred by flow is the concentration of some soluble factor that is released or absorbed by cells. An explanation consistent with the dependence of growth on serum concentration is that the signal is a serum-derived growth-promoting factor.

We thank Patricia Dopping-Hepenstal for growing 3T3 cells and Richard Adams for computer software.

Received 18 June; accepted 6 Sept 1984.

1. Dulbecco, R. *Nature* **227**, 802-806 (1970).
2. Rubin, H. & Rein, A. in *Growth Regulating Substances for Animal Cells in Culture* (eds Defendi, V. & Stoker, M.) 51-66 (Wistar Institute, Philadelphia, 1967).
3. Stoker, M. G. P. *Nature* **246**, 200-203 (1973).
4. Stoker, M. G. P. & Piggott, D. *Cell* **3**, 207-215 (1974).

5. Maroudas, N. G. *Cell* **3**, 217-219 (1974).
6. Whittenberger, B. & Glaser, L. *Nature* **272**, 821-822 (1978).
7. Dewey, C. F., Bussolari, S. R., Gimbrone, M. A. & Davies, P. F. *Trans. Am. Soc. mech. Engrs* **103**, 177-185 (1981).
8. Franke, R.-P., Grafe, M., Schnittler, H., Seiffge, D. & Mittermayer, C. *Nature* **307**, 648-649 (1984).
9. Daily, J. W. & Nece, R. E. *Trans. Am. Soc. mech. Engrs* **D82**, 217-232 (1960).

Immunoglobulin-like nature of the α -chain of a human T-cell antigen/MHC receptor

Charles H. Hannum*, John W. Kappler*, Ian S. Trowbridge†, Philippa Marrack* & John H. Freed*

* Division of Basic Immunology, Department of Medicine, National Jewish Hospital and Research Center, 3800 E. Colfax Avenue, Denver, Colorado 80206, USA

† The Salk Institute, La Jolla, California 92037, USA

Although the receptor with which T cells bind specific antigen can, like immunoglobulin, distinguish between antigens which differ only slightly in structure, it is unique in recognizing antigen only in conjunction with one of the self proteins of the major histocompatibility complex (MHC restriction)¹⁻⁴. The receptor was identified and characterized in mouse and man by using monoclonal antibodies to receptor idiotypes⁵⁻⁷, and consists of two disulphide-linked polypeptides, an acidic α -chain and a neutral to slightly basic β -chain^{5,8,9}. Peptide maps have shown that, like immunoglobulin, both chains vary for receptors of different specificities¹⁰⁻¹². T-cell-derived cDNA clones have recently been identified in mouse and man encoding immunoglobulin-like molecules¹³⁻¹⁵. These were identified as derived from β -chain genes through a partial N-terminal protein sequence of the β -chain isolated from a human T-cell tumour¹⁶. We have now purified the α - and β -chains of the receptor of the human T-cell leukaemia line HPB-MLT, and have determined the amino acid sequence of several tryptic peptides derived from each chain. Our results further confirm that the previously reported cDNA clones encode β -chains. The sequence of the α -chain peptides identify this as another immunoglobulin-like polypeptide chain. Particularly striking was an α -chain peptide with high homology to the conserved portion of the immunoglobulin J segment and T-cell receptor β -chains. Surprisingly, the α -chain peptides show little similarity to the sequence predicted by two overlapping putative murine α -chain cDNA clones¹⁷.

Using a monoclonal antibody to the receptor on the HPB-MLT cell line¹⁰, we purified several nanomoles of receptor trace-labelled with ¹²⁵I. The α - and β -chains of the receptor were separated and purified by two-dimensional non-reduced/reduced SDS-polyacrylamide gel electrophoresis (PAGE) as described in Fig. 1 legend. Both chains were cleaved

with trypsin and the resultant peptides were purified by reverse-phase HPLC (Fig. 2). Two β -chain peptides, designated Y and Z, were sequenced (Fig. 3). Identical sequences are predicted by the human cDNA sequence of Yanagi *et al.*¹³, further confirming this cDNA as encoding a β -chain and increasing our confidence that the other chain we had purified on the basis of its covalent linkage to this β -chain was indeed a *bona fide* α -chain.

Eight α -chain peptides were sequenced (labelled A-H in Fig. 2a); except for a few amino acids and in some cases the terminal lysine or arginine, an unambiguous sequence was obtained for each peptide. These sequences, listed in Fig. 3, represent the first reported for an α -chain. Each of the peptides was compared with all known sequences in the Dayhoff protein data base¹⁸ using the Unitary Matrix and the Mutation Data Matrix comparison programs. Obviously, because the peptides are so short each peptide showed sequence similarities to a different random set of bacterial, viral and eukaryotic proteins, but in no case was a perfect match found. However, most importantly, only homologies to immunoglobulins and other T-cell receptor chains occurred among the peptides.

The most striking similarity was found for peptide 'E', which shows clear homology to both T-cell receptor β -chain and immunoglobulin J segments (Fig. 4). The strongest homology is 75% (6/8) with the J _{β 1-2} segment from the β -chain of the human T-cell tumour Molt-3 (refs 13, 19) and with the murine T-cell β -chain J _{β 2-3} (ref. 20). Note that these homologies occur towards the C-terminal end of the J segments, within the strongly conserved region. This great similarity to β -chain J segments suggests that peptide E might be a particularly high yield peptide derived from some β -chain contaminating the purified α -chain. This possibility is eliminated by the absence of the peptide in the tryptic digest of HPB-MLT β -chain (Fig. 2). The identification of a J segment in the α -chain not only identifies its immunoglobulin-like nature but also predicts that the α -chain is encoded by separate V, J and C genes, and that immunoglobulin-like V-J rearrangements must occur in T cells to produce a functional α -chain gene. Confirmation of this prediction must await isolation of this gene.

Peptides F, G and H also show some similarity to immunoglobulins and other T-cell receptor sequences. Although these similarities are less striking than that of peptide E, they do strengthen the conclusion that the α -chain is an immunoglobulin-like molecule. Peptide G is worth discussing in some detail. This proline-rich peptide contains an acid-cleavable bond between aspartic acid and proline (Fig. 3). We have identified a single acid cleavage site in the HPB-MLT-derived α -chain roughly in the middle of the molecule (C.H.H., unpublished data). By analogy with immunoglobulin and other T-cell

Fig. 1 Purification of the α - and β -chains of the receptor of HPB-MLT. Lysate was prepared as described previously^{7,8,10} from 5×10^{10} HPB-MLT cells and combined with lysate prepared from 5×10^8 125 I-surface labelled HPB-MLT cells. The lysate was incubated overnight with excess Sepharose 4B beads coupled with the anti-receptor monoclonal antibody T40/25. The bound receptor was eluted with 0.05 M diethylamine, pH 11, and immediately neutralized with Tris buffer. A portion of the eluate corresponding to 5×10^8 cell equivalents was boiled in Laemmli sample buffer²⁴ and run in a single lane of a 10% SDS-PAGE slab gel without reduction. The lane was sliced from the gel, soaked briefly in sample buffer containing 5% 2-mercaptoethanol, laid along the top of a second 10% slab gel and electrophoresed again. The gel was stained with Coomassie blue (a) and an autoradiogram of the dried gel prepared (b). The 46,000-molecular weight (46K) α -chain and 38K β -chain are indicated in each panel, with molecular weight standards positioned along the diagonal containing proteins whose molecular weight did not change on reduction. For preparation of α - and β -chain from the rest of the eluate, this analytical method was scaled up. One-fifth of the T40/25 Sepharose eluate was loaded along the entire top of a 10% slab gel and run without reduction, with cytochrome c added as carrier to protect the receptor from oxidation. The gel was sliced into thin horizontal strips and those containing receptor (~ 85 K), detected by radioactivity, were soaked briefly in sample buffer containing 5% 2-mercaptoethanol, loaded along the top of a second slab gel, and electrophoresed again including 0.1 mM thioglycollate in the cathodic buffer chamber²⁵. This gel was also cut into thin horizontal slices and those corresponding to the α - and β -chains were identified by radioactivity. Finally, all slices containing either α - or β -chain were combined and pulverized. The proteins were eluted with water, dialysed extensively, reduced again with dithiothreitol, alkylated with 4-vinylpyridine²⁶, concentrated by lyophilization and precipitated with acetonitrile. The overall yield from five purifications was ~ 2 -3 nmol of each chain.

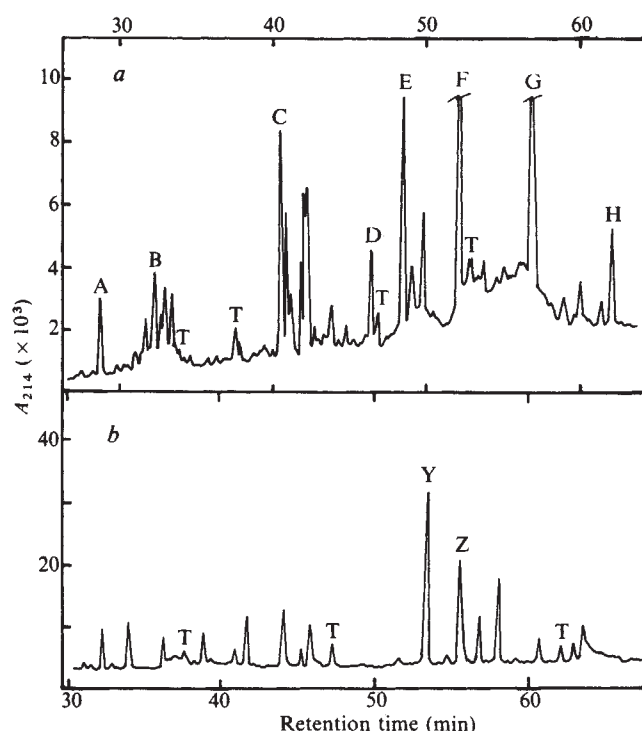
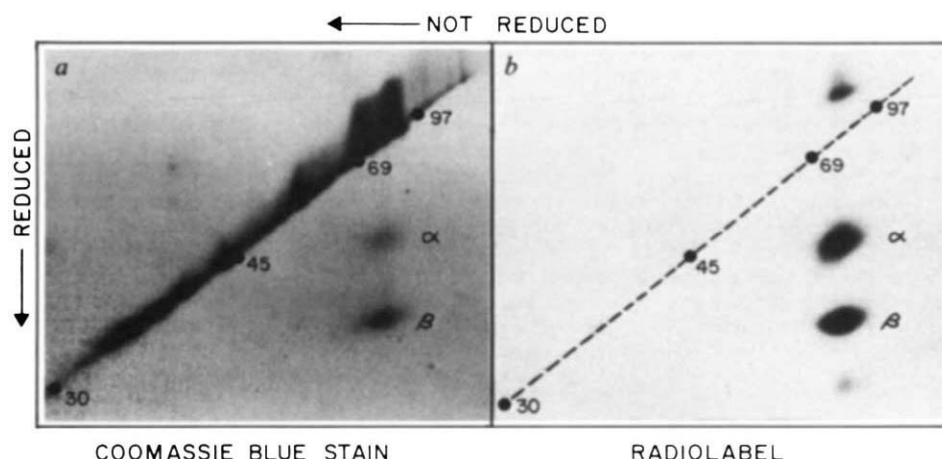


Fig. 2 Preparation of tryptic peptides from α - and β -chain. *a*, The elution pattern of the α -chain peptides. The peptides selected for sequencing are labelled A-H. *b*, The elution pattern of the β -chain peptides. The peptides selected for sequencing are labelled Y and Z. In both panels peptides known to be derived from the autodigestion of trypsin are labelled T.

Methods: Samples containing 1-2 nmol of α -chain and 2-3 nmol of β -chain were resuspended in 0.5 ml of 0.05 M ammonium bicarbonate and treated with 10% (w/w) TPCK-trypsin added as two equal aliquots over a 16-h incubation at 37 °C. The reaction was stopped by lyophilization. The lyophilized pellet was resuspended in 100 μ l of 0.1% trifluoroacetic acid (TFA) in water (solvent A). The insoluble core peptides were removed by centrifugation at 2,000g for 20 min. The soluble peptides were injected on a 4.6 $\times$ 250 mm Altex Ultrasphere-I.P. 5 μ C-18 reverse-phase column. The column was eluted at 1 ml min⁻¹, with an 80-min linear gradient of 0-50% solvent B (0.1% TFA in acetonitrile) in solvent A²⁷. The column effluent was monitored at 214 nm and 0.5-ml fractions collected.

receptor proteins, this would place the peptide towards the N-terminal end of the constant region. Similar proline-rich stretches occur in the beginning of the constant region of T-cell receptor β -chains^{13,15} and in many immunoglobulin light- and heavy-chain constant regions²¹. In immunoglobulins, these occur in the β -pleated sheets involved in the interaction between light and heavy chains²², and it is tempting to propose a similar structure and function in T-cell receptor α - and β -chains.

One of the most surprising results of our study is the relatively slight similarity of the α -chain peptides of the HPB-MLT receptor to the reported sequence of a proposed murine α -chain predicted from the pHDS4/203 cDNA clones isolated from a

cytotoxic T-cell clone¹⁷. Overall, the HPB-MLT peptides are no more similar to this sequence than they are to immunoglobulins or T-cell receptor β -chains. Given the extraordinary homology between mouse and human β -chains, it seems unlikely that the function of the T-cell receptor could tolerate such divergence between mouse and human in α -chain sequence. Rather, one must consider the possibility that this murine cDNA sequence does not encode a receptor α -chain, but some other immunoglobulin-like molecule. Although this might seem unlikely, several properties of this protein are inconsistent with the known properties of murine α -chains. Murine α -chains are acidic proteins, with N-linked carbohydrate adding as much as

	Position																	
Peptide	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
α chain																		
A	Leu	Val	Glu	Lys														
B	()	Asn	Ser	Ala	Val	Ala	Trp	Ser	(Ala)	Lys								
C	Asp	Ser	Asp	Val	Tyr	Ile	Thr	Asp	(Lys)									
D	Thr	Val	Leu	Asp	Met	Arg												
E	Ile	Ile	Phe	Gly	Ser	Gly	Thr	Arg										
F	Tyr	Ser	Trp	Asn	Phe	Gln	(Lys)											
G	Leu	Ser	Ile	()	Pro	Asn	Ile	Gln	Asn	Pro	Asp	Pro	Ala	Val	Tyr	Gln	Leu	(Arg)
H	Glu	Asp	Val	Glu	Gln	Ser	Leu	Phe	Leu	Ser	Val	()						
β chain																		
Y	Val	Ser	Ala	Thr	Phe	Trp	Gln	Asn	Pro	()								
Z	Ala	Lys	Pro	Val	Thr	Gln	Ile	Val	Ser	Ala	Glu	Ala	Trp	Gly	()			

Fig. 3 Amino acid sequence of the α - and β -tryptic peptides. Fractions from the HPLC, containing individual peptides, were pooled and evaporated to dryness by lyophilization. Sequences were determined on an Applied Biosystems gas-phase sequencer using either the '01' or '02' chemistry with TFA conversion of the thiazolinones to phenylthiohydantoins (PTH). Analysis of the PTH-derivatives produced was carried out on a Waters Novapak C-18 reverse-phase column eluted with a complex gradient of 16% acetonitrile in 40 mM sodium acetate, pH 5.0 (solvent A) and 60% isopropyl alcohol in water (solvent B). PTH-derivatives were monitored at 254 nm. Residues in parentheses indicate uncertain assignments. Open parentheses indicate that no assignment could be made at that position, although when at the end of a peptide the open parentheses indicate either a lysine or arginine.

Human	Peptide E	I	I	F	G	S	G	T	R
	J β 1-2	Y	T	F	G	S	G	T	R
	J β 1-1	A	F	F	G	Q	G	T	R
	J β 1-1	V	F	F	G	K	G	T	R
	J β 1-3	L	Y	F	G	E	G	S	R
	J β 1-4	L	F	F	G	H	G	T	K
	J β 1-5	P	L	F	G	E	G	T	R
	J β 1-6	L	Y	F	A	A	G	T	R
	J β 2-1	Q	F	F	G	P	G	T	R
	J β 2-2	L	Y	F	G	E	G	S	K
	J β 2-3	L	Y	F	G	S	G	T	R
	J β 2-4	L	Y	F	G	A	G	T	R
	J β 2-5	Q	Y	F	G	P	G	T	R
	J β 2-7	Q	Y	F	G	P	G	T	R
Human	J pHDS4/203	K	V	F	A	E	G	T	K
	J β consensus	-	-	W	G	Q	G	T	L
	J α consensus	-	-	F	G	Q	G	T	K
	J λ consensus	-	-	F	G	G	G	T	K

Fig. 4 Comparison of the sequence of peptide E of the α -chain with that of J segments of other T-cell receptor chains and of immunoglobulins. The sequence of peptide E of the receptor α -chain of HPB-MLT is shown aligned with portions of the J segment of human and mouse β -chains^{13,19,20,28}, the J segment predicted by the pHDS4/203 cDNA clones¹⁷, and consensus J segments for human immunoglobulins²¹. Identical residues are boxed. The standard one-letter amino acid code is used: A, alanine; R, arginine; N, asparagine; D, aspartic acid; C, cysteine; Q, glutamine; E, glutamic acid; G, glycine; H, histidine; I, isoleucine; L, leucine; K, lysine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine.

10,000–15,000 to their molecular weight^{5,8,9,23}. The pHDS4/203 cDNA encodes a basic protein with no sites for N-glycosylation. Furthermore, peptide maps of receptor α -chains have indicated variation between T-cell clones of different specificities^{10–12}. However, the expression of the genes detected by pHDS4/203 seems to involve only one V region and one J segment. Further

experiments will be necessary to resolve these apparent discrepancies.

We thank Craig R. Miles for technical assistance in obtaining the N-terminal amino acid sequences, Jim Leibson, Ella Kushnir and Janice White for other valuable technical help, Dr Ralph Kubo for help in the preparation of the immunosorbent beads and for sharing his protein chemistry experience with us, Dr George Tarr of the University of Michigan for his advice in our early attempts to sequence these proteins, and Kelly Bakke, Ruth Roberts and Edna Squillante for secretarial help. This work was supported by USPHS grants CA-36700 (to J.H.F.) and AI-17134 (to J.W.K. and P.M.) and American Cancer Society grants IM-49 (to J.W.K. and P.M.) and FRA-218 (to J.W.K.).

Received 3 September; accepted 27 September 1984.

1. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701–702 (1974).
2. Bevan, M. *Nature* **269**, 417–419 (1977).
3. Zinkernagel, R. M. *et al. J. exp. Med.* **147**, 882–896 (1978).
4. Kappler, J. & Marrack, P. *Nature* **262**, 797–798 (1976).
5. Allison, J. P., McIntyre, B. W. & Bloch, D. J. *Immun.* **129**, 2293–2300 (1982).
6. Meuer, S. *et al. J. exp. Med.* **157**, 705–719 (1983).
7. Haskins, K. *et al. J. exp. Med.* **157**, 1149–1169 (1983).
8. Kappler, J. *et al. Cell* **34**, 727–737 (1983).
9. Meuer, S. C. *et al. Nature* **303**, 808–810 (1983).
10. Kappler, J. *et al. Cell* **35**, 295–302 (1983).
11. McIntyre, B. W. & Allison, J. P. *Cell* **34**, 739–746 (1983).
12. Acuto, O., Meuer, S. C., Hodgdon, J. C., Schlossman, S. F. & Reinherz, E. L. *J. exp. Med.* **158**, 1368–1373 (1983).
13. Yanagi, Y. *et al. Nature* **308**, 145–149 (1984).
14. Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149–153 (1984).
15. Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153–158 (1984).
16. Acuto, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 3851–3855 (1984).
17. Saito, H. *et al. Nature* **309**, 757–762 (1984).
18. Barker, W. C. *et al. Protein Sequence Database of the Protein Identification Resource*, Release 2 (National Biomedical Research Foundation, Washington DC, 1984).
19. Siu, G. *et al. Cell* **37**, 393–401 (1984).
20. Gascoigne, N. R. J., Chien, Y.-H., Becker, D. M., Kavalier, J. & Davis, M. M. *Nature* **310**, 387–391 (1984).
21. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. in *Proteins of Immunological Interest* (National Institutes of Health, Bethesda, 1983).
22. Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961–997 (1979).
23. McIntyre, B. W. & Allison, J. P. *Cell* (in the press).
24. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
25. Hunkapiller, M. W., Lujan, E., Ostrander, F. & Hood, L. E. *Meth. Enzym.* **91**, 227–236 (1983).
26. Inglis, A. S. *Meth. Enzym.* **91**, 26–36 (1983).
27. Meek, J. L. & Rossetti, Z. L. *J. Chromat.* **211**, 15–28 (1981).
28. Chien, Y.-H., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322–326 (1984).

Cell-surface antigens expressed on L-cells transfected with whole DNA from non-expressing and expressing cells

Charles Hsu, Paula Kavathas & Leonard A. Herzenberg

Department of Genetics, Stanford University, Stanford, California 94305, USA

We have shown previously¹ that transfection of mouse L-cells with DNA from JM, a human T-cell line expressing certain T-cell differentiation antigens, yields stable transfectants expressing one or another of these antigens. The identities of the antigens were confirmed by immunoprecipitation and SDS-polyacrylamide gel electrophoresis. We now report that our procedure—co-transfection² with the chicken thymidine kinase gene (*tk*) and whole cellular DNA, selection with hypoxanthine-aminopterin-thymidine (HAT)³, and staining of the cells with fluorochrome-conjugated monoclonal antibodies and fluorescence-activated cell-sorter (FACS) selection—yields transfectants expressing a variety of cell-surface molecules (19 of 21 investigated), most at a frequency of about one per 10³ Tk⁺ transformants. Of these, 9 of 12 were transferred and expressed as readily using DNA from cells which did not express the cell-surface antigens as from tissues or cells that did express them.

We performed Tk co-transfections with DNA from human placenta, kidney, a renal cell carcinoma derived from the same individual as the kidney, and LCL-721, a human B-lymphoblastoid cell line. None of these tissues expresses serologically detectable Leu-1 or Leu-2, except LCL-721, which expresses Leu-1; in none is Leu-2 messenger RNA detectable by Northern blotting using cloned Leu-2 complementary DNA as probe²⁸. DNA from JM, which expresses these genes, served as a positive control.

Tk⁺ transformants were selected by FACS for expression of Leu-1 or Leu-2. As positive controls, we selected transfectants expressing HLA class I antigens or human transferrin receptor, as these are constitutively expressed in all of the tissues we studied. We recovered Leu-2⁺ transfectants from all of the

Table 2 Murine lymphoid antigens detected on transfected L-cells using monoclonal antibodies

Antigen	Tissue source of DNA			L(Tk ⁻)
	Spleen	Kidney	Muscle	
Ly-1	3/3*	2/2	2/2	5/6
Ly-2	2/3	2/2	2/2	4/6
BLA-2	2/3	2/2	2/2	3/6
ThB	3/3	2/2	2/2	ND
FcR(ref. 26)	2/3	0/2	1/2	ND
Thy-1	1/3	2/2	1/2	ND
L3T4a(ref. 27)	1/3	ND	1/2	4/6
H-2 ^d	3/3	2/2	ND	ND

* No. of dishes yielding positive transfectants/total no. of dishes analysed.

human DNA sources, at a frequency of about one per 10³ Tk⁺ transformants (Table 1). We also recovered, though less frequently, transfectants expressing either Leu-1 (from both antigen-expressing and non-expressing sources) or human transferrin receptor.

In our original work with JM DNA, one of the Leu-2⁺ transfectants expressed very high, but unstable, levels of surface Leu-2. The gene for Leu-2 was highly amplified in this transformant^{4,28}. We have now found that 11 of 20 independent Leu-2⁺ transfectants exhibit similarly high and heterogeneous levels of Leu-2. In 10 of these cases, further rounds of selection for the most brightly staining cells have yielded at least a 4-fold increase in antigen expression, resulting in cells that are now over 20-fold brighter than 'normal', stable Leu-2⁺ transfectants. In some cases, we have confirmed Leu-2 gene amplification by quantitative filter hybridization of RNA⁵ and DNA⁶ to a cloned Leu-2 gene probe²⁸.

To further our study using multiple organ and tissue sources and to determine how many of the antigens for which we had available antibodies could be transferred and expressed, we extended our work to mouse DNA transfections. We studied 11 murine cell-surface antigens. Nine are believed to be monomeric glycoproteins. One, Lyt-2, occurs primarily as a heterodimer⁷. The two chains may be encoded by one gene or

Table 1 Human antigens detected on transfected mouse L-cells using monoclonal antibodies

Antigen	DNA donor				
	Placenta	Kidney	Renal cell carcinoma	LCL-721 (B-lymphoblastoid)	JM (T-cell lymphoma)
HLA (W6/32)	ND	3/6†	6/6	4/4	2/2
Leu 2	6/6	3/6	1/6	7/8	3/6
(Leu 2)*	(4/6)	(1/3)	(1/1)	(3/7)	(2/3)

Isolation of transfectants expressing novel cell-surface antigens: L(Tk⁻) cells²² were grown in 100-mm tissue culture dishes (Lux) in Dulbecco's modified Eagle's medium with 10% fetal calf serum. DNA was prepared from various tissues as described previously²³. Briefly, the tissues were pulverized under liquid nitrogen, dissolved in 1% sarkosyl, 50 mM EDTA, 0.2 mg ml⁻¹ proteinase K (Merck), followed by treatment with RNase, extraction with phenol followed by chloroform, and caesium chloride equilibrium gradient centrifugation. The purified DNA was recovered (and sterilized) by ethanol precipitation and dissolved in 10 mM Tris-HCl, 1 mM EDTA. Transfections were carried out using a modification of the protocol of Wigler *et al.*²⁴. In most cases the DNA was sheared before precipitation by one passage through a 23-gauge needle. In all cases, 1 µg of Tk-bearing plasmid (the chicken thymidine kinase gene, cloned into the plasmid pBR322 and propagated in the host *Escherichia coli* GM48 (*dam*⁻*mec*⁻)²⁵ was used with 20 µg of genomic ('carrier') DNA per dish of 10⁶ cells. Transfectants were selected for expression of Tk genes by growth in HAT³. They were then selected for expression of specific antigens by staining with fluoresceinated monoclonal antibodies and FACS sorting, as described elsewhere¹. Initial dishes of transfectants were pooled, if necessary, to form dishes of 1,000–1,500 Tk⁺ clones each. These were then collected, and each dish was divided into aliquots of 10⁶ cells, one aliquot for each reagent to be screened. Staining was done at nominally saturating concentrations of antibody, in sterile 96-well V-bottom culture trays. In the initial round of sorting, cells falling within arbitrary 'gates' representing the positively staining cells were sorted into pots and then regrown for further analysis. In some cases it proved possible to clone positively staining cells directly from the initial culture. In most cases it was necessary to enrich with one or more rounds of 'pot-sorting' for bright cells before it became feasible to clone true positives into individual wells.

W6/32 is an anti-HLA framework monoclonal antibody that is specific for HLA heavy chains even when they are in association with mouse β₂-microglobulin^{20,21}. ND, not determined.

* Amplified phenotype (no. amplified/no. positive).

† No. of dishes yielding positive transfectants/total no. of dishes analysed.

Table 3 Expression of mouse genes in transfected L-cells requires mouse DNA

DNA source	Human Leu-2a	Human NGF receptor	Antigen Rat NGF receptor	Mouse 30-E2	Mouse L3T4a
PC12(rat)	ND	ND	5/8*	0/8	0/8
A875(human)	4/6	1/6	ND	0/6	0/6
Peripheral blood lymphocytes (human)	6/6	ND	ND	0/6	0/6

PC12 is a rat chromocytoma expressing NGF receptor; A875 is a human melanoma expressing NGF receptor but not Leu-2. Transfectant cultures were subjected to at least four rounds of sorting and regrowth before being typed as negative. In previous experiments with 30-E2 and L3T4a, no more than two rounds of FACS selection had ever been required to obtain cultures that were >90% enriched for positive cells, when there were positives present in the starting dishes of transfectants. We note that the results with A875 provide another example of efficient transfer of Leu-2 using DNA from a non-expressing source.

* No. of dishes yielding positive transfectants/total no. of dishes analysed.

by closely linked genes. In the case of H-2, the host cell⁸ or the culture medium serum⁹ can supply β_2 -microglobulin to complement the heavy chains encoded by the donor DNA, and permit surface expression.

We used DNA from BALB/c mice, as they differ at several serologically detectable loci from the mouse strain (C3H) from which L-cells were derived¹⁰. Because all of the antigens tested are present on at least 10% of spleen cells^{11,12}, we used spleen DNA as our positive standard. None of these antigens is known to be expressed on any of the other tissues used. The various DNA sources differed little in yield of transfectants for 8 of the 11 antigens tested (Table 2). For the other three antigens (T200, BLA-1 and 30-Fl, not tabulated), we recovered only one transfectant, preventing meaningful comparisons.

For heuristic purposes, we also used DNA from the L(Tk⁻) line itself. We used only those antibodies with which we had previously most readily detected transfectants. We recovered L-cell transfectants positive for the target antigens as frequently with L-cell DNA as with the other DNAs (Table 2).

It is conceivable that the murine antigens are produced by L-cell genes that are active in rare variants in normal L-cell cultures, or which are activated by the transfection process. We therefore transfected L-cells with human or rat DNA and looked for cells expressing murine antigens. We did not isolate any transfectants expressing the murine antigens tested, although we did obtain transfectants expressing Leu-2 or rat nerve growth factor (NGF) receptor at expected frequencies (Table 3). Furthermore, in the case of Ly-1, Lyt-2 and H-2 antigens, the BALB/c DNA-derived transfectants stain with antibodies specific for the donor allotype, showing directly that they express transferred, and not endogenous, genes.

Two antigens for which we obtained transfectants were Lyt-2 (see Table 2) and 30-Fl (data not shown). In the normal animal, each is associated with another cell-surface determinant (Lyt-2 with Lyt-3 (ref. 11) and 30-Fl with BLA-1 (R. Hardy, unpublished)). None of the Lyt-2⁺ transfectants stain with anti-Lyt-3, nor do they contain the Lyt-3 polypeptide when examined by immunoprecipitation and SDS-polyacrylamide gel electrophoresis. However, the one 30-Fl⁺ transfectant isolated also expresses BLA-1. As co-transfer of two unlinked genes is very unlikely¹³, and in light of their association on normal cells, we propose that 30-Fl and BLA-1 are encoded by one gene or by two very tightly linked genes.

The BLA-2⁺ transfectants expressed extremely high levels of antigen, enabling us to perform an initial characterization of the molecule. We immunoprecipitated a 94,000 molecular weight protein consisting of one polypeptide chain. This molecule has now been identified on various murine B-cell lines known from staining studies to be BLA-2⁺ (R. Hardy and G. Calicchio, unpublished).

We believe this to be the first demonstration that several genes can be stably transferred in active form using DNA from tissue sources not expressing those genes. Further, in the case of most of the genes we analysed, DNA from non-expressing sources

appears to be as efficient in transfection as DNA from expressing tissues.

Other studies¹⁴⁻¹⁶, using cloned genes, show an inverse correlation between the activity of genes in gene transfer and their degree of cytosine base methylation. The means by which the genes in our study were originally turned off may not depend on this type of modification, or they may be insufficiently modified to keep them turned off in L-cells. We have readily recovered Leu-2⁺ transfectants using the highly methylated¹⁷ DNA from human sperm. However, we have not recovered any Leu-2⁺ transfectants using DNA from certain human choriocarcinoma cell lines (S. Alberti, C.H. and L.A.H., in preparation); these are also the only DNA sources we have used which do not express antigens of the major histocompatibility complex. We are now investigating whether there is a special mechanism for turning off genes in these cells.

Enhancer sequences have been shown to be important in the expression of certain cloned genes in transfection systems^{18,19}. These *cis*-acting elements increase transcription rate and accuracy, and seem to act in a highly tissue-specific manner. The stringent tissue-specificity attributed to the expression of certain differentially expressed genes (for example, immunoglobulins) in gene-transfer systems does not seem to hold in L-cells for the genes we have transferred using total cellular DNA.

We thank Dr R. R. Hardy for purifying and fluorescein-conjugating most of the anti-mouse antibodies used in this work, and Drs F. Fitch, J. Unkeless, L. Lanier, P. Jones and E. Shooter for gifts of reagents. This work was supported in part by NIH grants CA-04681, GM-28428 and GM-17367.

Received 8 May; accepted 21 August 1984.

- Kavathas, P. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 524-528 (1983).
- Wigler, M. *et al.* *Cell* **16**, 777-785 (1979).
- Szybalska, E. H. & Szybalski, W. *Proc. natn. Acad. Sci. U.S.A.* **48**, 2026-2034 (1962).
- Kavathas, P. & Herzenberg, L. A. *Nature* **306**, 385-387 (1983).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
- Southern, E. J. *molec. Biol.* **98**, 503-517 (1975).
- Ledbetter, J. A., Seaman, W. E., Tsu, T. T. & Herzenberg, L. A. *J. exp. Med.* **153**, 1503-1516 (1981).
- Woodward, J. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **79**, 3613-3617 (1982).
- Bernabeu, C., van de Rijn, M., Lerch, P. G. & Terhorst, C. P. *Nature* **308**, 642-645 (1984).
- Altman, P. A. & Katz, D. D. (eds) *Inbred and Genetically Defined Strains of Laboratory Animals* FASEB (1979).
- Ledbetter, J. A. & Herzenberg, L. A. *Immun. Rev.* **47**, 63-90 (1979).
- Hardy, R. R., Hayakawa, K., Haaijman, J. & Herzenberg, L. A. *Nature* **297**, 589-591 (1982).
- Warrick, H., Hsiung, N., Shows, T. B. & Kucherlapati, R. J. *Cell Biol.* **86**, 341-346 (1980).
- Christy, B. & Scangos, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6299-6303 (1982).
- Stein, R., Razin, A. & Cedar, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3418-3422 (1982).
- Busslinger, M., Hurst, J. & Flavell, R. A. *Cell* **34**, 197-206 (1983).
- Jagiello, G., Tantravahi, U., Fang, J. S. & Erlanger, B. F. *Exptl Cell Res.* **141**, 253-259 (1982).
- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-728 (1983).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-740 (1983).
- Barnstable, C. J. *et al.* *Cell* **14**, 9-20 (1978).
- Brody, F. & Parham, P. J. *Immun.* **128**, 129-135 (1982).
- Kit, S., Dubbs, D., Piekarski, L. & Hsu, T. *Exptl Cell Res.* **31**, 291-312 (1963).
- Blattner, F. R. *et al.* *Science* **202**, 1279 (1978).
- Wigler, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 1373-1376 (1979).
- Marinus, M. G. *Molec. gen. Genet.* **127**, 47-55 (1973).
- Mellman, I. S. & Unkeless, J. C. *J. exp. Med.* **152**, 1048-1069 (1980).
- Dianylas, D. P. *et al.* *Immun. Rev.* **74**, 29-56 (1983).
- Kavathas, P., Herzenberg, L. A., Sukhatme, U. & Parnes, J. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Evidence for two distinct *c-src* loci on human chromosomes 1 and 20

Michelle M. Le Beau, Carol A. Westbrook,
Manuel O. Diaz & Janet D. Rowley

Section of Hematology/Oncology, Department of Medicine,
University of Chicago, Chicago, Illinois 60637, USA

A number of proto-oncogenes have recently been localized to the chromosomal segments that are the breakpoints in the specific rearrangements noted in human malignant diseases¹. Moreover, rearranged forms of several proto-oncogenes have been identified in malignant cells; in several instances, the proto-oncogene has undergone an alteration as a result of a nonrandom chromosomal rearrangement²⁻⁴. One proto-oncogene that has yet to be associated with human neoplastic disease is *c-src*, the cellular homologue of the transforming sequence of Rous sarcoma virus (RSV). By somatic cell hybridization, *c-src* has been mapped to chromosome 20, but its precise location was not determined⁵. We have now mapped this gene by using *in situ* hybridization of the cloned human *c-src* probe to human mitotic chromosomes. We report here that the human genome contains two loci with strong homology to the coding regions of this oncogene, at 1p34-p36 and 20q12-q13. It is noteworthy that these chromosomal regions are frequently involved in the structural rearrangements observed in haematological malignant diseases⁶⁻⁸.

In situ hybridization⁹ of a ³H-labelled *c-src* probe containing genomic *c-src* sequences spanning exons 10-12 (plasmid pHu-1-c) resulted in specific labelling of two chromosomal regions, the distal short arm of chromosome 1 and the long arm of chromosome 20. Figure 1 shows the distribution of labelled sites in 100

metaphase cells examined. Of these cells, 40 (40%) were labelled on region 1p3, bands p34-p36, of one or both chromosomes 1; these sites represented 13.5% (68 of 505) of all labelled sites ($P < 0.005$). On chromosome 20, 34 metaphase cells (34%) exhibited label on the mid- to terminal segment of 20q. Fifty labelled sites, or 9.9% of all labelled sites, were noted on 20q12-q13 ($P < 0.005$). The distribution of grains from labelled chromosome 1 ($n = 65$) and labelled chromosome 20 ($n = 51$) indicates significant clustering of grains to bands 1p34-p36 and to bands 20q12-q13 (Fig. 2). A small cluster of grains was observed at band 9q34 ($P < 0.005$, 1.8% of labelled sites), suggesting that cross-hybridization to *c-abl* had occurred¹⁰. When the stringency of hybridization was increased (hybridized slides washed extensively at 42 °C rather than 39 °C), labelling was not observed in this region. However, specific labelling of 1p34-p36 (9.6% of labelled sites) and 20q12-q13 (7.9% of labelled sites) was maintained under more stringent conditions (60 metaphase cells examined). *In situ* hybridization with an 800-base pair *PvuII* fragment containing *v-src*-specific sequences derived from the Schmidt-Ruppin strain of RSV (plasmid pPvuIIIE)¹¹ confirmed the presence of these two loci (data not shown). Although the signal was much weaker than that observed with the human *c-src* probe, a clustering of grains was again observed on 1p34-p36 and 20q12-q13.

RSV is one of the most highly oncogenic of the known retroviruses. Infection of avian fibroblasts with RSV results in a high degree of malignant transformation; this transforming capability of RSV is mediated by the *v-src* sequences¹². Interestingly, infection of cultured mouse bone marrow with RSV results in increases in stem cell proliferation and in self-renewal, two properties associated with myeloproliferative disorders¹³.

We have now regionally mapped the proto-oncogene *c-src* to two sites in the normal human karyotype by using *in situ* hybridization. The fact that the two loci hybridized to both the

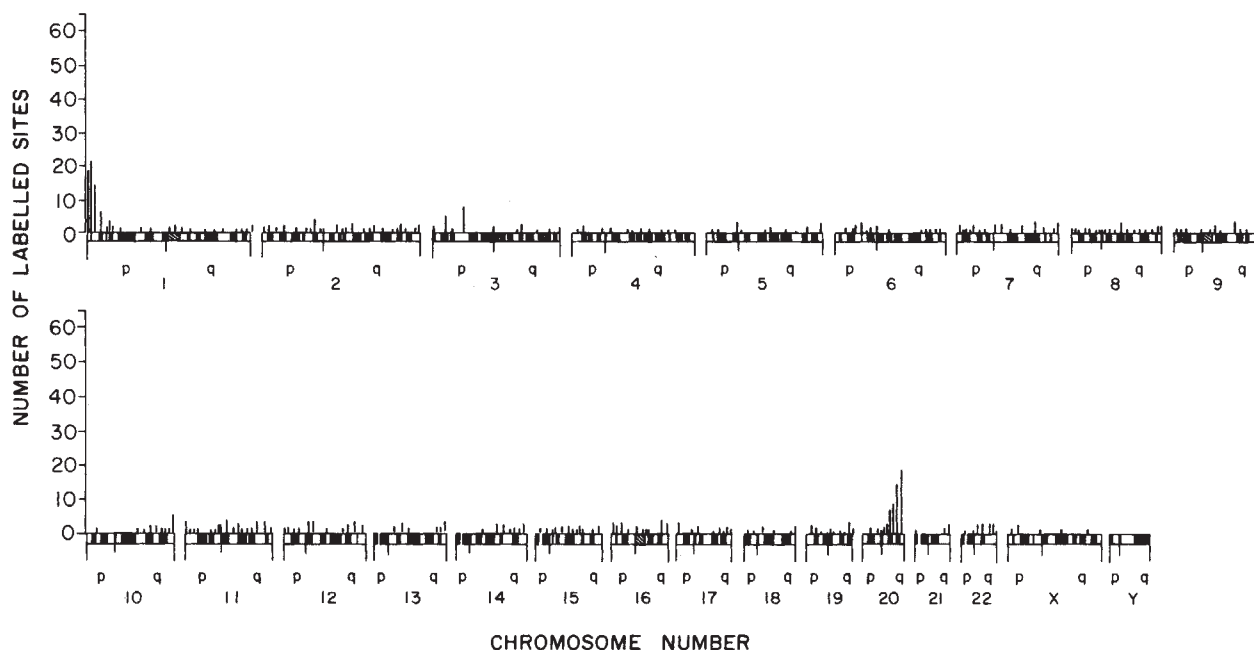


Fig. 1 Sub-localization of the human *c-src* gene to two human chromosomes. Distribution of labelled sites in 100 metaphase cells identified significant clustering of grains on regions p3 and q1 of chromosomes 1 and 20, respectively. Human metaphase cells prepared from phytohaemagglutinin-stimulated peripheral blood lymphocytes were hybridized with the ³H-labelled pHu-1-c plasmid DNA which contained a 1.7-kilobase *c-src*-specific insert, encompassing exons 10, 11 and 12 of the human genomic *src* sequence, cloned in pUC8. Radiolabelled probes were prepared by nick translation of the entire plasmid using all four ³H-labelled dNTP to specific activity 5×10^7 – 1×10^8 d.p.m. per μ g. Metaphase cells were hybridized at 10 or 30 ng probe per μ l for 16 h at 37 °C, exposed to autoradiographic emulsion for 8–15 days, and G-banded with the Wright's staining method as described by Harper⁹. The following modifications of the method of Harper were used. Chromosomal DNA was denatured by immersion of slides in 70% formamide in $0.6 \times$ SSC (0.09 M NaCl/0.009 M trisodium citrate) at 70 °C, pH 7.0. Radiolabelled probe was hybridized in 50% formamide/ $2 \times$ SSCP (0.3 M NaCl/0.03 M trisodium citrate/0.04 M NaPO₄, pH 5.0)/10% dextran sulphate/3 mM Tris-0.3 mM EDTA (pH 7.0). After hybridization, the slides were washed extensively in 50% formamide/ $1 \times$ SSC, 39 °C, pH 7.0, followed by washes in $2 \times$ SSC and by ethanol dehydration. For improved chromosome banding, autoradiographs were destained with 95% ethanol-HCl (0.25%, 10 M HCl) and restained, either once or twice.

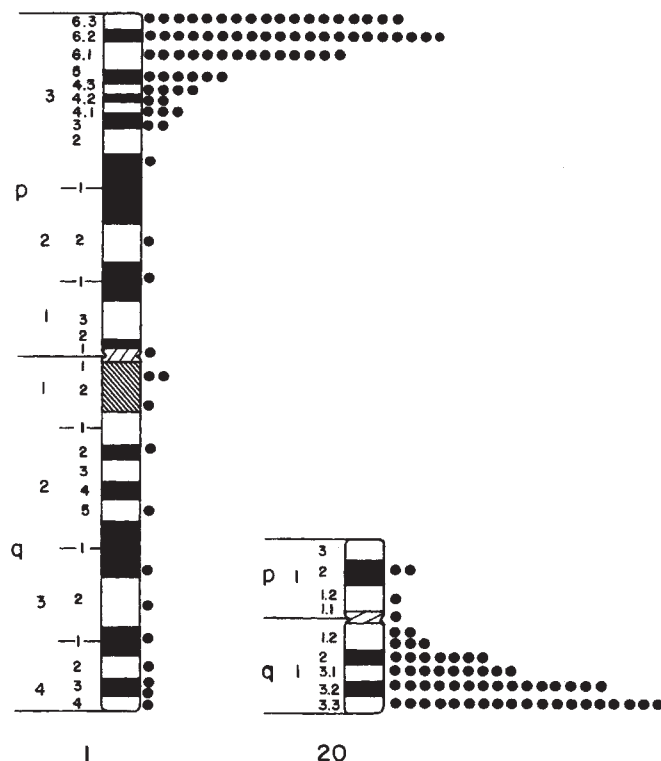


Fig. 2 Distribution of labelled sites on 65 and 51 labelled chromosomes 1 and 20, respectively. Of 86 total sites on chromosome 1, 68 (79%) were located on 1p34-pter. Of 59 labelled sites on chromosome 20, 50 (84.7%) were noted on q12-qter.

v-src and *c-src* probes suggests that the homology lies within the coding regions of the genes rather than in the intervening sequences that are absent in the *v-src* probe. With man-mouse somatic cell hybrids, sequences homologous to *v-src* had previously been mapped to chromosome 20⁵ and, recently, to its long arm (A. Y. Sakaguchi, personal communication). Southern blot analysis of less strongly hybridizing bands that did not segregate with chromosome 20⁵, suggested that additional loci may exist. Moreover, two related DNA clones corresponding to two separate *src* genes have been isolated from a human genomic library (R. Parker, G. Mardin, R. Lebo, H. Varmus and J. M. Bishop, in preparation), and were assigned to chromosomes 1 and 20 by hybridization of specific probes to dot blots of human chromosomes isolated by fluorescence activated sorting. Whether both of these genes are transcribed is not known; however, human cells contain two related, but distinct, protein forms of pp60^{c-src} (ref. 14).

Translocations involving bands p34-p36 of chromosome 1 have been noted frequently in haematological neoplasia and in solid tumours^{7,8}. We have observed rearrangements involving these bands in 5% of patients with acute non-lymphocytic leukaemia or a myelodysplastic syndrome, and in 10% of patients with non-Hodgkin's lymphoma (E. Brown, J.D.R. and M. M. Le B., unpublished results). Nonrandom deletions of 1p with a breakpoint between 1p13 and 1p36 have been identified in neuroblastomas¹⁵ and in malignant melanomas¹⁶. Abnormalities of chromosome 20 have been observed in a number of patients with haematological malignant diseases. Although loss of one chromosome 20 and translocations involving 20q have been reported, a more common abnormality is a deletion of the long arm of the chromosome. Such a deletion has been noted in patients with acute non-lymphocytic leukaemia or myelodysplastic syndrome, as well as in those with polycythaemia vera^{6-8,17,18}. This abnormality has been described as a terminal deletion involving band q11 (ref. 18). However, at the light

microscopy level, we cannot exclude the possibility that the deletion is actually an interstitial one with breakpoints at q11 and q13. *In situ* hybridization of *c-src* to malignant cells with a del(20q) would provide a definitive answer. We are now investigating these possibilities, along with an analysis of cells with rearrangements of chromosome 1. In view of our findings that *src*-specific sequences are localized to two human chromosomes that are frequently involved in rearrangements, a role for *c-src* in the pathogenesis of human neoplastic diseases should be considered. In this regard *c-src*-related sequences have been found to be expressed in malignant cells of two patients with haematopoietic neoplasms¹⁹.

We thank J. Michael Bishop for the gift of the cloned *src* probes and for valuable discussion, Diane Roulston for technical assistance, and Nanci Carrothers for help in the preparation of the manuscript. This work was supported in part by US Department of Energy, contract DE-AC02-80EV 10360, PHS grant CA 16910 from the National Cancer Institute, by the University of Chicago Cancer Research Foundation (J.D.R., M.D.), ACS grant IN-41-W-14 (M.D.), and by a Leukemia Research Foundation Grant (M.D.). M.M.L. is a Fellow of the Leukemia Society of America. C.A.W. is a Fellow of the Damon Runyon-Walter Winchell Cancer Foundation.

Received 6 June; accepted 16 August 1984.

- Rowley, J. D. *Nature* **301**, 290-291 (1983).
- Leder, P. *et al. Science* **222**, 765-771 (1983).
- ar-Rushdi, A. *et al. Science* **222**, 390-393 (1983).
- Bishop, J. M. *Cell* **32**, 1018-1020 (1983).
- Sakaguchi, A. Y., Naylor, S. L. & Shows, T. B. *Prog. Nucleic Acid Res. molec. Biol.* **29**, 279-283 (1983).
- Rowley, J. D. & Testa, J. R. *Adv. Cancer Res.* **36**, 103-148 (1982).
- Mitelman, F. *Cytogenet. Cell Genet.* **36**, 1-516 (1983).
- Fourth International Workshop on Chromosomes in Leukemia *Cancer Genet. Cytogenet.* **11**, 249-360 (1984).
- Harper, M. E. & Saunders, G. F. *Chromosoma* **83**, 431-439 (1981).
- Heisterkamp, N. *et al. Nature* **299**, 747-749 (1982).
- De Lorge, W. J., Lucin, P. A., Goodman, H. M., Varmus, H. E. & Bishop, J. M. *J. Virol.* **36**, 50-61 (1980).
- Hanafusa, H. *Compreh. Virol.* **10**, 401-483 (1977).
- Boettiger, D., Anderson, S. & Dexter, T. M. *Cell* **36**, 763-773 (1984).
- Shealy, D. J. & Erikson, R. L. *Nature* **293**, 666 (1981).
- Brodeur, G. M. *et al. Cancer Res.* **41**, 4678-4686 (1981).
- Balaban, G. *et al. Cancer Genet. Cytogenet.* (in the press).
- Testa, J. K., Kinnealey, A., Rowley, J. D., Golde, D. W. & Potter, D. W. *Blood* **52**, 868-877 (1978).
- Davis, M. P., Dewald, G. W., Pierre, R. V. & Hoagland, H. C. *Cancer Genet. Cytogenet.* **12**, 63-71 (1984).
- Slamon, D. J., de Kernion, J. B., Verma, J. & Cline, M. J. *Science* **224**, 256-262 (1984).

Biological properties of human *c-Ha-ras1* genes mutated at codon 12

Peter H. Seeburg, Wendy W. Colby,
Daniel J. Capon, David V. Goeddel
& Arthur D. Levinson

Department of Molecular Biology, Genentech, Inc.,
460 Point San Bruno Boulevard, South San Francisco,
California 94080, USA

Vertebrate genomes contain proto-oncogenes whose enhanced expression or alteration by mutation seems to be involved in the development of naturally occurring tumours^{1,2}. These activated genes, usually assayed by their ability to induce the malignant transformation of NIH 3T3 cells, are frequently related to the *ras* oncogene³⁻¹⁰ of Harvey (*Ha-ras*)¹¹ or Kirsten (*Ki-ras*)¹² murine sarcoma viruses, or a third member of this family (*N-ras*)¹³⁻¹⁵. Activation involves point mutation which often affect codon 12 (refs 16-26) of the encoded 21,000-molecular weight polypeptide (p21). To provide insight into structural requirements involved in p21 activation, we have now constructed 20 mutant *c-Ha-ras1* genes by *in vitro* mutagenesis, each encoding a different amino acid at codon 12. Analysis of rat fibroblasts transfected with these altered genes demonstrates that all amino acids except glycine

(which is encoded by normal cellular *ras* genes) and proline at position 12 activate p21, suggesting a requirement for an α -helical structure in this region of the polypeptide. The morphological phenotype of cells transformed by the activated genes can, however, depend on the particular amino acid at this position.

The human c-Ha-*ras*1 gene is contained on a genomic 4.7-kilobase (kb) *Bam*HI-*Sph*I DNA fragment¹⁷ and its coding region is distributed across four exons¹⁹. To generate codon 12 mutants, a DNA fragment encompassing exon 1 was cloned into bacteriophage M13 vector mp10 (ref. 27) and exhaustive mutagenesis at codon 12 was performed using techniques similar to those described by Zoller and Smith²⁸. The oligodeoxynucleotides for introducing the desired mutations were synthesized²⁹ in three pools, each pool containing oligomers designed to alter the sequence of codon 12 by one, two or three nucleotides, respectively (Fig. 1). The few mutants not identified in this manner were obtained by repeating the mutagenesis and screening with individually synthesized oligomers.

All mutations so generated were inserted by *in vitro* recombination back into the c-Ha-*ras*1 gene (Fig. 2). This strategy was used to minimize the possibility of obtaining transformed cell clones due to the significant overexpression of otherwise non-transforming genes which can occur when strong, heterologous promoters are used³⁰. To assess the transformation potential of the 20 c-Ha-*ras*1 plasmids, each was individually transfected onto contact-inhibited Rat-1 cells³¹ by calcium phosphate-mediated DNA transfection in the presence of the Tn5-encoded phosphotransferase (*neo*) gene^{32,33}; cells incorporating the exogenous DNA were selected in the presence of the antibiotic G418 (refs 32, 33). Cotransfected markers usually become incorporated together into recipient cells^{34,35} such that individual G418-resistant colonies could be examined for the expression of the Ha-*ras* genes independently of the morphological phenotype of the cells. In this manner we could directly confirm that cells co-transfected with non-transforming c-Ha-*ras*1 genes indeed expressed the encoded polypeptide.

When co-transfected with most of the mutant *ras* genes, 80–90% of the G418-resistant colonies contained cells which

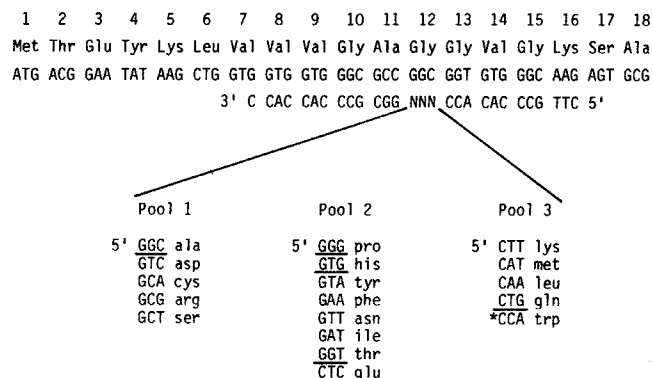


Fig. 1 Mutagenesis of c-Ha-*ras*1 codon 12. The first 18 codons of the proto-oncogene are shown, as carried on the recombinant M13 phage mp10/*ras*-Gly 12. For mutagenesis, three oligomer pools of the sequence 5'-CTTGCCACACCCNNNGGCCGCCA-CCACC-3' were synthesized by a phosphotriester approach²⁹. The actual nucleotides occupying the NNN positions of the three pools are shown; these were designed to alter the codon 12 sequence by one, two or three nucleotides, respectively. To reduce the complexity of pool 2, the oligomer changing glycine to tryptophan was included in pool 3 (see asterisk).

Methods: Oligomers (in pools or individually) were phosphorylated using α -ATP and polynucleotide kinase, and annealed to single-stranded mp10/*ras*-Gly DNA. Annealings (10 μ l; 10 mM Tris-HCl pH 7.5, 7 mM MgCl₂, 0.1 mM EDTA, 50 mM NaCl) containing 0.1 pmol of template DNA and 0.1 pmol of phosphorylated oligomer were incubated for 5 min at 55 °C. *In vitro* DNA synthesis was initiated by adding 40 μ l of 50 mM Tris-HCl pH 7.5, 7 mM MgCl₂, 0.1 mM EDTA, 1 mM ATP, 100 μ M each of dATP, dGTP, dTTP and dCTP, 5 U *E. coli* DNA polymerase I, large fragment (Boehringer Mannheim) and 400 U T4 DNA ligase (New England Biolabs). After incubation for 3 h at 37 °C, reaction mixtures were extracted with phenol, DNA was precipitated with EtOH, and 2–4% of the DNA was used to transform *E. coli* strain JM103 made competent for DNA uptake. Transformation mixtures (300 μ l) were spread onto bacterial plates (150 mm), yielding several hundred phage plaques. To screen for mutants, these plates were probed by *in situ* hybridization on duplicate nitrocellulose filter-lifts⁴⁵. One set of filters was probed with the oligonucleotide 5'-CTTGCCACACCCGCCGCCA-3', fully complementary to the original sequence (coding for glycine) and other with a mutant-specific oligonucleotide or oligomer pool. Oligonucleotides were phosphorylated to a specific activity of $\sim 10^7$ d.p.m. per pmol with [γ -³²P]ATP (ICN, crude >4,000 Ci mmol⁻¹) and polynucleotide kinase (NEN). Hybridization was overnight using 10⁷ c.p.m. per filter. To screen for the original sequence, annealings were in 5 $\times$ SSC, 45% formamide, 50 °C and excess radioactivity was washed off in 2 $\times$ SSC, 0.1% SDS at room temperature, followed by 15 min at 50 °C. Conditions for detecting mutants were as follows. For mutants derived from pool 1, a set of 13-mers was used as a probe (5'-NNNGGCCGCCACC-3', with NNN represented by GGC, GTC, GCG and GCT). This oligomer mixture was hybridized at 20 °C in 5 $\times$ SSC, 30% formamide, and filters were washed in 2 $\times$ SSC, 0.1% SDS at the same temperature. To probe for mutants derived from pools 2 and 3, oligomer mixtures from these pools were hybridized in 5 $\times$ SSC, 40% formamide at 60 °C and filters were washed in 2 $\times$ SSC, 0.1% SDS at 20 °C followed by a 30-min incubation at 30 °C. All filters were blotted dry after the last wash and exposed to X-ray film for several hours at -70 °C using an intensifying screen. For those mutants not obtained by using pools 1, 2 and 3 (see underlined codons), the following oligomers were synthesized individually and served as mutagenesis primers as well as mutant-specific probes:

5'-CTTGCCACACACGCGGCCACC-3' (Ala)
GGG (Pro)
GTG (His)
GGT (Thr)
CTG (Gln)

To probe for mutants, hybridization was in 5 $\times$ SSC at 20 °C and filters were washed in 5 $\times$ SSC at 20 °C, incubated in 3 M trimethylammonium chloride at 55 °C (W. Wood, unpublished) and briefly rinsed in 2 $\times$ SSC at 20 °C. Phage identified by hybridization as carrying an altered 12th codon were grown in small cultures (2 ml), and DNA was isolated and characterized by dideoxy sequencing^{46–48}.

Table 1 Transformation properties of *ras* mutants

Position 12 codon	Activated genes	Focus formation	Agar growth
Gly	c-Ha- <i>ras</i> 1 (refs 16–19), c-Ki- <i>ras</i> 2 (ref. 41), N- <i>ras</i> (ref. 15)	—	—
Ile		+++	+
Val	T24/EJ Ha- <i>ras</i> 1 (refs 16–19), SW480 Ki- <i>ras</i> 2 (ref. 21)	+++	+
Leu		+++	+
Ser	v-Ki- <i>ras</i> (ref. 12)	++	+
Thr		+++	+
Ala		++	+
Met		++	+
Cys	CaLu Ki- <i>ras</i> 2 (refs 20, 21) PR371 Ki- <i>ras</i> 2 (ref. 23)	++	+
Tyr		++	+
Phe		++	+
Trp		++	+
Pro		—	—
Lys	v-BALB- <i>ras</i> (ref. 17)	+	+
Arg	v-Ha- <i>ras</i> (ref. 11), v-Ra- <i>ras</i> (ref. 26), A1698, A2182 Ki- <i>ras</i> 2 (ref. 24)	+++	+
His		++	+
Asp	134-51 Ha- <i>ras</i> 1 (ref. 22)	++	+
Glu	NMU-Ha- <i>ras</i> (ref. 25)	++	+
Asn		++	+
Gln		+	+

Focus formation was determined by the ability of *ras*-cotransfected Rat-1 cells to form foci (see Fig. 3 legend). Agar growth was assessed by plating individual *ras*-cotransfected subclones, or cotransfected mass cultures, into soft agar⁵¹ at two densities (4×10^3 and 2×10^4 per 60 mm dish) and scoring colonies of >30 cells after 2–3 weeks.

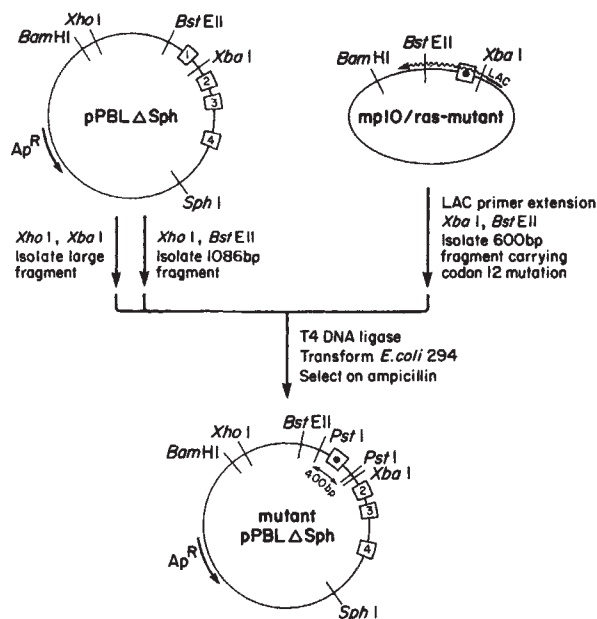


Fig. 2 Construction of *c-Ha-ras1* genes substituted at codon 12. Plasmid *pPBLΔSph*, a derivative of *pPBL116-4* (ref. 19) lacking the 1.7-kb *Sph*I-*Bam*HI region (containing repeated elements)¹⁹ of the 6.5-kb *Bam*HI fragment, was cleaved with endonucleases *Xho*I and *Xba*I. The larger fragment (6.7 kb) containing sequences of the bacterial plasmid pBR322 and *ras* exons 2, 3 and 4, together with the presumed transcriptional terminator of the *c-Ha-ras1* gene, was isolated from a 6% polyacrylamide gel. The 1,086-bp *Xho*I-*Bst*EII fragment containing the prelude region of *c-Ha-ras1* (ref. 19) was similarly isolated. DNA carrying the first exon of *c-Ha-ras1* mutated in codon 12 was obtained by gel isolation of a 600-bp *Bst*EII-*Xba*I fragment cleaved from *in vitro* synthesized DNA after priming with the 17-mer LAC primer⁴⁶ on the *mp10/ras* mutant phage DNAs. The *Bst*EII-*Xba*I fragment with the valine codon was isolated from plasmid pT24-10 (ref. 19). A three-component ligation with these DNA fragments yielded the mutant variants of plasmid *pPBLΔSph* used for transfections. To ensure that each mutant plasmid carried the correction codon substitution, plasmids were screened for the absence of a *Nae*I restriction site spanning codon 12 (refs 16, 17), and a 400-bp *Pst*I fragment containing exon 1 was isolated from each plasmid DNA and analysed by DNA sequencing.

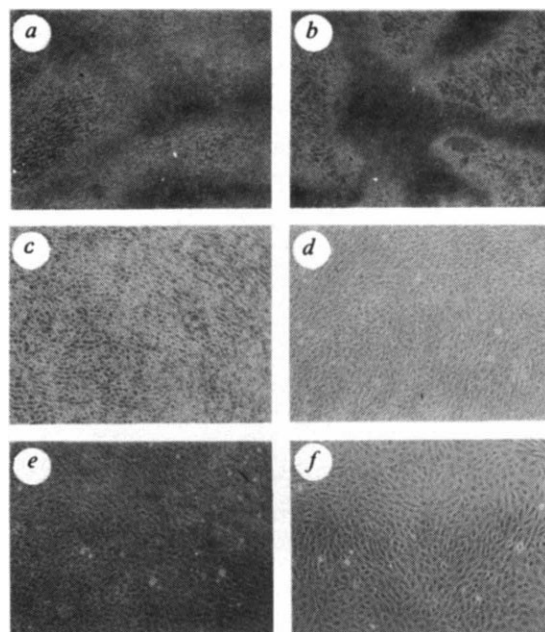


Fig. 3 Phase-contrast photomicrographs of Rat-1 cells ($\times 40$). Cells were transfected by *c-Ha-ras1* genes encoding substitutions at codon 12 of: a, valine; b, isoleucine; c, tyrosine; d, lysine; e, proline; f, untransfected Rat-1 cells.

Methods: Transfections were performed by a modification of the calcium phosphate precipitation method⁴⁹ on 1.0×10^6 Rat-1 cells using 5 μ g of each of the *c-Ha-ras1*-containing plasmids and 500 ng of pSVNeoBal6, a plasmid containing the neomycin-resistance^{32,33} gene isolated from Tn5 (ref. 33) under transcriptional control of the simian virus 40 early promoter (a 348-bp *Pvu*II-*Hind*III fragment⁵⁰). After 48 h, the transfected cells were split 1:3 into selective medium containing Dulbecco's modified Eagle's medium with 10% fetal bovine serum (Hyclone) and 400 μ g ml⁻¹ G418. After 14 days, resistant colonies were cloned using cloning cylinders, or alternatively on occasion passed *en masse*. Monolayers derived from individual clones or mass cultures were scored for morphology and photographs taken 8-16 days after the cells reached confluency. Foci which overgrew the monolayers were scored for their morphological phenotype (indicated by + to +++ on Table 1) as described in the text.

were not contact-inhibited, unlike control cells which received only the *neo*-resistance plasmid (see Table 1). The morphological phenotype induced by these activated *ras* genes was reproducibly dependent upon the particular mutation involved. Mutations which gave rise to gene products with the most potent transforming ability (measured by colony morphology) were those encoding valine, leucine, isoleucine, arginine and threonine at position 12. Cells expressing such gene products typically displayed a fully transformed morphology: the cells were consistently round and refractile, and they grew to the highest saturation densities. An intermediate class of mutants (serine, methionine, cysteine, tyrosine, phenylalanine, tryptophan, histidine, aspartate, glutamate, alanine and asparagine at position 12) generated transformants that overgrew the monolayer, although the cells displayed a less striking change in morphology. Mutants with the most subtle transforming activity (lysine and glutamine at position 12) induced foci only after extended times (typically 2-3 weeks). Such cells exhibited a nearly normal morphology. We failed to observe any focus formation in cells co-transfected with *ras* genes encoding either glycine or proline at position 12 throughout five separate experiments, each performed in duplicate. Figure 3 shows representative morphologies of G418-resistant Rat-1 cells co-transfected with these four classes of mutants.

To provide an additional measure of the transformed phenotype, we determined whether cultures co-transfected with each mutant *ras* vector could grow in an anchorage-independent

fashion. We found that all G418-resistant cell lines capable of forming foci on monolayers grew in soft agar (Table 1) with plating efficiencies of 0.2-6.0% (data not shown). Colonies co-transfected with Gly 12 and Pro 12 versions of *c-Ha-ras1* genes were unable to grow in agar, even when mass cultures derived from many (10-30) G418-resistant colonies were plated. Despite the additional selection for anchorage-independent growth following G418 selection, cells which were originally flat (those expressing Lys 12 or Gln 12 versions of p21) did not display any obvious morphological changes when replated following passage through agar (data not shown).

These results indicate that the presence of either a glycine or proline residue at position 12 of p21 results in a polypeptide incapable of inducing the morphological transformation of Rat-1 cells, at least when expressed under the transcriptional control of its own promoter. To verify that these vectors indeed directed the synthesis of p21, as well as to determine the relative levels of expression of p21, cloned cells from cultures co-transfected with representative mutant *ras* genes were examined for the expression of the *c-Ha-ras1* p21 gene product. Cellular proteins were labelled with ³⁵S-methionine and extracts were immunoprecipitated with rabbit immune sera raised against the *c-Ha-ras1* gene product expressed in *Escherichia coli*³⁶. The immunoprecipitates were resolved by SDS-polyacrylamide gel electrophoresis and visualized by fluorography. Figure 4a illustrates that anti-p21 antiserum detects the *c-Ha-ras1* gene product in G418-resistant colonies co-selected with active transforming

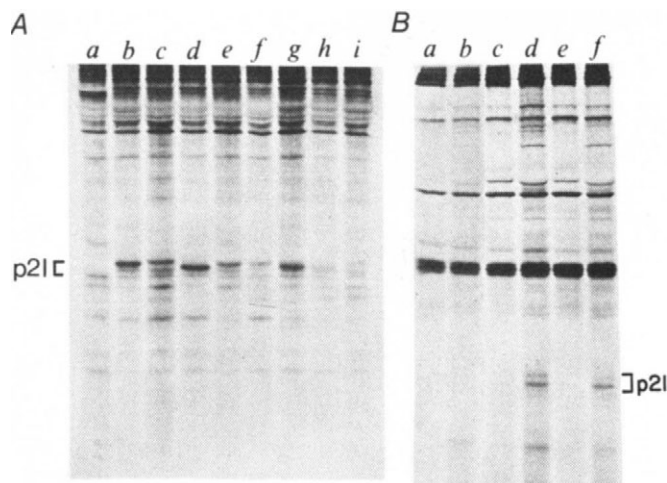


Fig. 4 Expression of p21 in transfected cell lines. **A**, ^{35}S -methionine-labelled cell lines were immunoprecipitated with rabbit anti-p21 serum (lanes b–i) or with preimmune serum (lane a). Cells were transfected by c-Ha-ras1 genes substituted at position 12 by: lanes a, b, valine; lane c, isoleucine; lane d, threonine; lane e, tyrosine; lane f, phenylalanine; lane g, arginine; lane h, lysine; lane i, untransfected Rat-1 cells. **B**, lysates were immunoprecipitated by anti-p21 rabbit serum (lanes b, d, f) or with preimmune rabbit serum (lanes a, c, e). Lanes a, b, untransfected Rat-1 cells; lanes c, d, cells were transfected by pPBLΔSph (glycine); lanes e, f, cells were transfected by the Pro 12 mutant.

Methods: Cultures were labelled with $150\ \mu\text{Ci}\ ^{35}\text{S}$ -methionine (Amersham) per ml methionine-free media for 3 h. Cells were lysed in buffer containing 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.5% sodium deoxycholate, 0.5% Nonidet P-40, 0.1% SDS, and 0.5% Aprotinin (Sigma). Lysates were preincubated with formaldehyde-fixed *Staphylococcus aureus* (Pansorbin, Calbiochem) for 30 min, and the bacteria were removed by centrifugation. The supernatants were immunoprecipitated with the anti-p21 polyclonal rabbit serum³⁶ or preimmune serum for 1 h at 4°C. Immune complexes were precipitated by *S. aureus* and the bacteria were washed three times with the lysis buffer. Immunoprecipitated proteins were analysed by electrophoresis on 15% or 12.5% SDS-polyacrylamide gels⁵² and visualized by fluorography after treatment with Enhance (NEN)

genes. Notably, each *ras* polypeptide exhibits a slightly different electrophoretic mobility, extending previous descriptions of p21 polypeptides which display altered electrophoretic mobilities as a consequence of single amino acid substitutions^{37–39}. Human c-Ha-ras1 p21 was also detected in cell lines co-transfected with those c-Ha-ras1 genes unable to effect the morphological transformation of Rat-1 cells (Gly 12 and Pro 12) (Fig. 4b). In these cases, the p21(Gly 12) and p21(Pro 12) exhibit indistinguishable electrophoretic mobilities, as found for these polypeptides directly expressed in *E. coli* (W.W.C., unpublished results). These results verify that the failure of these two genes to transform Rat-1 cells was not due to unanticipated mutations preventing the expression of p21.

The observation that p21 *ras* can induce a panoply of morphological phenotypes in Rat-1 cells depending on the amino acid at position 12 is unexpected. Interestingly, certain replacements which engender p21 with a more restricted transformation potential *in vitro* (for example, Lys 12 and Ser 12) appear in retroviral cognates of transduced *ras* genes (v-bas (ref. 17) and v-Ki-ras (ref. 12) respectively). We suspect this underscores the contribution of quantitative effects as well as qualitative changes in p21 on transformation potential. Consistent with this, we have found that sequences extending 1.7 kb 3' of the *SphI* site in the c-Ha-ras1 gene (containing the reiterated copies of a 28 base pair (bp) element¹⁹) facilitate the expression of p21; when the expression of p21 (Lys 12) is augmented about fivefold in this manner, a more fully transformed phenotype is induced (W.W.C., unpublished results). These findings, together with the observation that extreme overexpression of even normal *ras* genes can induce the morphological transformation of NIH 3T3

cells³⁰, imply that a threshold level of expression exists for each position 12 *ras* mutant, above which any may transform such cells, although this remains to be established for p21(Pro 12).

Computer analysis of the secondary structure of p21 led to the suggestion that the amino-terminal domain of the polypeptide consists of two helical segments separated by a coil region surrounding amino acid 12 (ref. 22). It was predicted that, as substitution of any other amino acid for glycine at position 12 would probably extend the flanking helical regions²², any mutation affecting the coding properties of codon 12 would activate the oncogenic potential of p21 (refs 22, 40). We have tested this prediction and confirmed that while most amino acid substitutions can activate the polypeptide, those least compatible with an α -helical structure (glycine and proline⁴¹) encode polypeptides incapable of morphologically transforming Rat-1 cells, at least when expressed under the control of the normal *ras* promoter. We note, however, that beyond the marked tendency of glycine and proline residues to break α -helical regions within proteins, residues with the highest reported helical potential⁴¹ do not correlate in any obvious way with those at position 12 of p21 which induce the most pronounced changes in cellular morphology. Note that even some amino acids with strong transforming potential (isoleucine, threonine, leucine) have yet to be identified as activating lesions in tumours. Although it is possible that this reflects an inability of some of the 'activated' genes to participate in the oncogenic pathway by which *ras* genes exert their *in vivo* influence, it is most easily explained by the fact that such changes require the substitution of at least two of the three nucleotides which comprise codon 12, whereas all activated genes characterized to date vary from the cellular proto-oncogene by a single nucleotide. If indeed all but two amino acids at position 12 can activate p21 by *in vivo* criteria, it is possible that those with a potentially limited *in vitro* transforming capability escape easy detection when assayed by usual transfection methods; if this is so, the frequency of activated *ras* genes detected to date in tumours may in fact be an underestimate.

The functional alterations directly responsible for the oncogenic activation of human *ras* proteins are unclear. Attempts to discern differences between normal and activated *ras* polypeptides with regard to specificity or extent of nucleotide binding^{36,42,43}, acylation, or subcellular localization⁴⁴ have been unsuccessful. The observation that p21 possesses an inherent GTPase activity which is selectively impaired by an activating Val 12 mutation³⁶ raises the possibility that an inability to hydrolyse GTP may constitutively activate the polypeptide. The availability of *ras* mutants such as those described here and elsewhere³⁹ which encode polypeptides of varying transformation potentials should facilitate further efforts to define the biochemical changes critical for oncogenic activation of p21.

We thank Dr Charles Benton for the preparation of anti-p21 antisera and Joel Hayflick for assistance with the mutagenesis.

Received 21 June; accepted 21 August 1984.

1. Bishop, J. M. *Rev. Biochem.* **52**, 310–354 (1983).
2. Land, H., Parada, L. F. & Weinberg, R. A. *Science* **222**, 771–778 (1983).
3. Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
4. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
5. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. *Nature* **298**, 343–347 (1982).
6. Pulciani, S. *et al. Nature* **300**, 539–542 (1982).
7. McCoy, M. S. *et al. Nature* **302**, 79–81 (1983).
8. Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 2112–2116 (1983).
9. Eva, A., Tronick, S. R., Gol, R. A., Pierce, J. H. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4926–4930 (1983).
10. Fujita, J. *et al. Nature* **309**, 464–466 (1984).
11. Dhar, R. *et al. Science* **217**, 934–936 (1982).
12. Tsuchida, N., Ryder, T. & Ohtsubo, E. *Science* **217**, 937–939 (1982).
13. Hall, A., Marshall, C. J., Spurr, N. K. & Weiss, R. A. *Nature* **303**, 396–400 (1983).
14. Murray, M. *et al. Cell* **33**, 749–757 (1983).
15. Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. *Cell* **34**, 581–586 (1983).
16. Tabin, C. J. *et al. Nature* **300**, 143–149 (1982).
17. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
18. Taparowsky, E. *et al. Nature* **300**, 762–765 (1982).
19. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33–37 (1983).
20. Shimizu, K. *et al. Nature* **304**, 497–500 (1983).
21. Capon, D. J. *et al. Nature* **304**, 507–513 (1983).

22. Santos, E., Reddy, E. J., Pulciani, S., Feldman, R. J. & Barbacid, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4679-4683 (1983).
23. Nakano, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 71-75 (1984).
24. Santos, E. *et al. Science* **223**, 661-664 (1984).
25. Sukumar, S., Notario, V., Martin-Zanca, D. & Barbacid, M. *Nature* **306**, 658-661 (1983).
26. Rasheed, S., Norman, G. L. & Heidecker, G. *Science* **221**, 155-157 (1983).
27. Messing, J. & Vieira, J. *Gene* **19**, 264-276 (1982).
28. Zoller, M. J. & Smith, M. *Nucleic Acids Res.* **10**, 6487-6500 (1982).
29. Crea, R. & Horn, T. *Nucleic Acids Res.* **8**, 2331-2348 (1980).
30. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **297**, 479-483 (1982).
31. Freeman, A. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2415-2419 (1973).
32. Colbere-Garapin, F., Horodniceanu, F., Kourilsky, P. & Garapin, A.-C. *J. molec. Biol.* **150**, 1-14 (1981).
33. Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
34. Wigler, M. *et al. Cell* **16**, 777-785 (1979).
35. Perucho, M., Hanahan, D. & Wigler, M. *Cell* **22**, 309-317 (1980).
36. McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **310**, 644-649 (1984).
37. Der, C. J. & Cooper, G. M. *Cell* **32**, 201-208 (1983).
38. Feig, L. A., Bast, R. C., Knapp, R. C. & Cooper, G. M. *Science* **223**, 698-701 (1984).
39. Fasano, O. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4008-4012 (1984).
40. Pincus, M. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 5253-5257 (1983).
41. Chou, P. Y. & Fassman, G. D. *Biochemistry* **13**, 222-245 (1974).
42. Shih, T. Y. *et al. J. Virol.* **42**, 253-261 (1982).
43. Papageorge, A., Lowy, D. & Scolnick, E. J. *J. Virol.* **44**, 509-519 (1982).
44. Finkel, T., Der, C. J. & Cooper, G. M. *Cell* **37**, 151-158 (1984).
45. Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
46. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
47. Sanger, F. *et al. J. molec. Biol.* **143**, 161-178 (1980).
48. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
49. Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373-1376 (1979).
50. Toozé, J. (ed.) *DNA Tumor viruses: Molecular Biology of Tumor Viruses Pt 2* (Cold Spring Harbor Laboratory, New York, 1981).
51. Kaplan, P. L., Topp, W. C. & Ozanne, B. *Virology* **108**, 484-490 (1981).
52. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Growth-rate dependent regulation of mRNA stability in *Escherichia coli*

G. Nilsson*, J. G. Belasco†, S. N. Cohen† & A. von Gabain*

* Unit for Applied Cell and Molecular Biology, University of Umea, S-901 87 Umea, Sweden

† Departments of Genetics and Medicine, Stanford University School of Medicine, Stanford, California 94305, USA

The rate of production of bacterial gene products is known to vary with the rate of cell growth, the concentrations of many cellular proteins are altered during times of decreased growth rate¹. In addition, proteins whose *in vivo* levels show no significant alterations with changes in cell doubling time must be synthesized at rates that vary in direct proportion to the growth rate of the cell. In certain instances, growth-rate dependent gene regulation has been shown to occur at the transcriptional or translational level^{2,3}. Another potentially important element in the regulation of gene expression is the stability of messenger RNA. We report here the effect of bacterial growth rate on the half lives of four different monocistronic *Escherichia coli* mRNA species. The stabilities of two of these species, the transcripts of the *ompA* and *cat* genes^{4,5}, exhibit a marked dependence on cell growth rate, whereas the half lives of the transcripts of the *lpp* and *bla* genes^{7,8} are constant over a broad range of cell doubling times. Our results indicate that *E. coli* can alter the rate of synthesis of certain proteins by modulating mRNA stability in response to changes in the rate of cell growth.

The *ompA* gene of *E. coli*, which encodes a highly abundant outer membrane protein, specifies an mRNA transcript that has an unusually long half life^{4,5,9}. Previously, it has been shown that the amount of OmpA protein in the outer membrane remains constant throughout a wide range of bacterial doubling times¹⁰. Such invariability requires a mechanism for adjusting the rate of biosynthesis of this protein in direct proportion to the rate of cell growth. To determine whether this bioregulatory mechanism involves changes in the stability of the *ompA* transcript, we studied the decay of *ompA* mRNA in bacteria growing in conditions that yield a range of doubling times.

E. coli strain C600 was cultured at 30 °C in one of three bacterial growth media: L-broth, MOPS/glucose or MOPS/ace-

Table 1 Half lives of individual *E. coli* transcripts at different growth rates

	mRNA segment length (nucleotides)	Transcript half lives in bacteria grown in various media		
		L-broth (40 min)	MOPS/glucose (100 min)	MOPS/acetate (200 min)
<i>bla</i>	587	3.3 ± 0.4	3.2 ± 1.0	2.4 ± 0.7
	267	3.1 ± 0.1	3.2 ± 0.5	3.1 ± 0.3
	80	2.8 ± 0.9	2.5 ± 0.5	2.5 ± 0.1
<i>cat</i>		2.3 ± 0.6	—	0.4*
<i>lpp</i>		20 ± 1	—	23 ± 3
<i>ompA</i>	598	13 ± 3	6.8 ± 0.8	4.4 ± 0.6
	340	15 ± 2	5.0 ± 0.3	4.4 ± 0.4
	200	25 ± 2	10.6 ± 2.7	3.4 ± 0.6
	85	21 ± 4	5.9 ± 0.4	3.8 ± 1.0

Half lives are reported in minutes. Messenger RNA segment boundaries for the *ompA* and *bla* transcripts are as previously defined⁹. The cleavage of a single phosphodiester linkage within a segment of a transcript is scored as the degradation of that segment. All experiments were performed with *E. coli* strain C600 growing exponentially at 30 °C. **Methods:** Decay data for the *ompA* and *cat* transcripts were obtained as described in Figs 1 and 2. For the analysis of *ompA* mRNA decay in cells multiplying with a doubling time of 100 min, RNA was extracted from *E. coli* growing in MOPS/glucose¹ 0, 4, 8 and 16 min after rifampicin addition. The decay of the *lpp* transcript in bacteria growing in L-broth and MOPS/acetate was examined by RNA blotting⁹. The RNA samples were collected 0, 20, 40 and 60 min after rifampicin addition, fractionated by gel electrophoresis, blotted onto diazophenylthioether paper¹⁶, and probed with nick-translated plasmid pKEN113, which contains a copy of the *E. coli lpp* gene¹⁷. Segmental analysis of the degradation of *bla* mRNA in *E. coli* growing in L-broth, MOPS/glucose and MOPS/acetate was performed as previously described⁹. Half lives were determined by densitometric analysis of autoradiograms. Rates of decay were calculated by least-squares determination of the slope of a plot of $-\ln$ (band intensity) against time. Errors were estimated from the standard deviation of the slope. For the *bla* transcripts, the half lives of only three of the seven characteristic segments⁹ are reported.

tate¹; in these media, the cells doubled every 40, 100 or 200 min respectively. The decay of *ompA* mRNA, and of its component segments, after the blockage of transcription by rifampicin was analysed by S₁ nuclease protection as described previously⁹. Figure 1 and Table 1 show that the half life of the *ompA* transcript, as judged from the stability of its least stable segment, fell from about 15 min in cells cultivated in L-broth (doubling time, 40 min) to about 4 min in cells cultivated in MOPS/acetate (doubling time, 200 min). Nutritional conditions that led to an intermediate growth rate (MOPS/glucose; doubling time, 100 min) resulted in an intermediate stability for the *ompA* transcript. The retardation of decay reported previously for the segments of the *ompA* transcript nearest the 5' end⁹ was not observed at the slow growth rate. In MOPS/acetate, all segments of the transcript showed approximately the same rate of decay (Table 1). This degradation pattern exhibited by the *ompA* transcript in bacteria grown under conditions that lead to an extended doubling time presumably results from a failure of the cells to accumulate decay intermediates, as observed previously for *bla* mRNA. We have proposed that such a decay pattern reflects rate-limiting initial cleavage of the full-length transcript⁹.

These measurements of mRNA half life rely on the fact that almost all (>95%) of the total RNA in *E. coli* is represented by stable RNA species: that is, ribosomal RNA and transfer RNA¹¹. Because these abundant and stable RNA molecules do not decay to any significant extent during the time course of an experiment for determining mRNA half life¹², they provide a convenient and reliable reference for measuring the rate of degradation of labile transcripts such as mRNA.

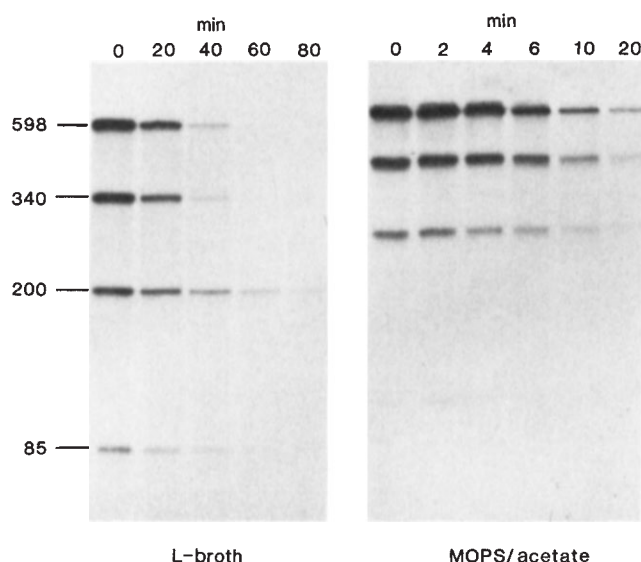


Fig. 1 Decay of *ompA* mRNA in *E. coli* cells growing with a doubling time of 40 min (left) or 200 min (right). Total cellular RNA from *E. coli* strain C600 growing exponentially at 30 °C in L-broth (doubling time, 40 min) was isolated 0, 20, 40, 60 and 80 min after rifampicin treatment⁹. Similarly, RNA from bacteria growing at 30 °C in MOPS/acetate¹ (doubling time, 200 min) was isolated 0, 2, 4, 6, 10 and 20 min after rifampicin addition. Segmental analysis of *ompA* mRNA decay was performed as previously described⁹. The bands correspond to the component segments of the *ompA* transcript; the order of the segments (5' to 3') is 200, 85, 598 and 340 nucleotides⁹.

In the absence of any compensatory effect on the initiation of transcription, a reduction in the stability of a transcript should result in a corresponding decrease in its intracellular concentration. Consistent with this prediction is our observation that the fraction of total cellular RNA represented by the *ompA* message falls by a factor of five in cells grown in MOPS/acetate as opposed to L-broth. We also find that the concentration of total cellular RNA, which consists primarily of the stable species rRNA and tRNA, drops by a factor of 1.8 when the bacterial doubling time increases from 40 min (cells grown in L-broth) to 200 min (cells grown in MOPS/acetate), in agreement with previous studies¹¹. Thus, we calculate that the steady-state concentration of *ompA* mRNA is reduced roughly by a factor of nine in bacterial cells growing in MOPS/acetate.

Because the concentration of OmpA protein in cells cultured under different growth conditions is constant¹⁰, the rate of synthesis of the OmpA protein must decrease by about a factor of five as the cell doubling time is increased from 40 to 200 min. The decreased rate of synthesis of OmpA protein that accompanies the prolonged cell doubling time in MOPS/acetate can be accounted for entirely by the observed change in *ompA* mRNA concentration; moreover, most of the reduction in the concentration of the *ompA* transcript can be attributed to its decreased stability.

To determine whether the observed shortening of the *ompA* mRNA half life in cells growing in MOPS/acetate medium is the result of slowed growth rate *per se*, or whether the effect is otherwise related to the media used, bacteria were grown in L-broth until their doubling time began to decrease on entry into the stationary phase. At a stage of growth corresponding to a doubling time of ~200 min, the cells were treated with rifampicin, and RNA was extracted at time intervals thereafter. Under these growth conditions, the half life of the *ompA* transcript was about 4 min, the same half life observed for *ompA* transcripts from cells having a 200-min doubling time resulting from growth in MOPS/acetate medium.

Like *ompA* mRNA, the *lpp* transcript encodes an abundant outer membrane protein (lipoprotein) and has the potential to form extensive secondary structure⁷. However, growth-rate

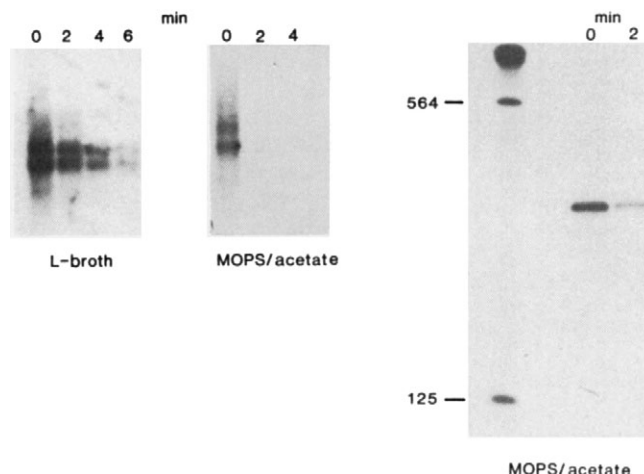


Fig. 2 Decay of *cat* mRNA in *E. coli* strain C600 containing the plasmid pACYC184 was grown to log phase at 30 °C in L-broth (left: doubling time, 40 min) or in MOPS/acetate (middle and right: doubling time, 200 min), and total cellular RNA was isolated 0, 2, 4 and 6 min after rifampicin treatment⁹. Analysis of *cat* mRNA decay was performed by fractionation of the RNA samples on a 1.2% agarose/formaldehyde gel²² followed by blotting onto a nitrocellulose filter²³ (left and middle). The blots were probed with nick-translated *cat* gene DNA. The decay of *cat* mRNA from *E. coli* growing in MOPS/acetate was also monitored by S₁ nuclease analysis²⁴ (right). Plasmid pACYC184 linearized with *Eco*RI and 5' end-labelled with [γ -³²P]ATP and polynucleotide kinase was used as the probe. The observed protected DNA fragment corresponds to the 5' terminal 270 nucleotides of *cat* mRNA²⁵. The finding of a unique 5' end by S₁ nuclease mapping demonstrates that the two *cat* messages detected by RNA blotting result from multiple sites of transcriptional termination.

dependent alteration of mRNA half life was not observed for the transcript of the *lpp* gene, which had a half life of ~22 min at 30 °C whether the bacteria were grown in L-broth or MOPS/acetate (Table 1). In addition, there was no change in the fraction of total RNA represented by the *lpp* transcript under these different growth conditions. Thus, an increased rate of mRNA decay with the slowing of cell growth is not a universal phenomenon, nor does it seem to be a general property of long-lived *E. coli* transcripts. Although it is not known whether the biosynthesis of the *lpp* gene product is stringently or loosely related to the rate of bacterial growth, our data demonstrate that any down-regulation of the rate of production of lipoprotein is not controlled at the level of mRNA stability.

The effect of growth rate on the stability of two *E. coli* mRNA species having the short half life more typical of *E. coli* messages was also examined. These were the *bla* transcripts of plasmid pBR322, which encode the periplasmic protein β -lactamase, and the *cat* transcripts of plasmid pACYC184, which encode the cytoplasmic protein chloramphenicol acetyltransferase (CAT)^{6,8,13-15}.

There are six overlapping *bla* messages transcribed from pBR322 (ref. 9). These transcripts are initiated at one of two promoters and are terminated at one of three sites. All six *bla* transcripts had a half life of 3 min independent of the growth medium used. Moreover, under all growth conditions, each of the component segments of *bla* mRNA⁹ decayed at the same rate (Table 1). When *bla* transcripts from cells isolated just before entry into the stationary phase of growth were analysed as described above for the *ompA* transcript, the half life was again found to be 2-3 min. As in the case of the *lpp* transcript, the fraction of total RNA represented by *bla* mRNA was the same at all growth rates, indicating that the steady-state concentration of *bla* mRNA fell by the same factor of 1.8 as did the overall RNA concentration in bacteria growing in MOPS/acetate. (The small decrease in *bla* mRNA concentration presumably reflects a slightly reduced rate of transcription.)

In contrast, the intracellular concentration of *cat* mRNA fell by a factor of ~13 when the cell doubling time was increased

Table 2 Enzyme activities of β -lactamase and chloramphenicol acetyltransferase from *E. coli* grown in different media

	L-broth	MOPS/acetate
β -Lactamase	13	43
CAT	0.37	0.25

Enzyme activities are units per mg of cellular protein. All experiments were performed with *E. coli* strain C600 containing either pBR322 or pACYC184. Different growth rates were achieved by growing cells in L-broth or in MOPS/acetate. Bacteria growing exponentially were ruptured by sonication, and the cell debris was removed by centrifugation. The supernatant was assayed for either β -lactamase activity or chloramphenicol acetyltransferase (CAT) activity^{18,19}. Total protein concentration was determined using the method of Lowry²⁰. The copy number²¹ of pACYC184 at the two growth rates was found to be the same, whereas the copy number of pBR322 increased twofold in MOPS/acetate.

from 40 to 200 min. Most of the reduction in *cat* mRNA concentration can be explained by the fact that the multiple transcripts of the *cat* gene, which result from the use of one promoter and two transcriptional terminators (Fig. 2), decayed approximately five times faster in response to the fivefold prolongation of doubling time. The *cat* messages isolated from bacteria cultivated in L-broth showed a half life of 2 min, whereas *cat* mRNA from cells grown in MOPS/acetate medium showed a half life of about 0.4 min (Table 1, Fig. 2).

Analysis of the activities of β -lactamase and CAT in cells growing in L-broth or MOPS/acetate shows that the concentrations of both enzymes, relative to total cellular protein, are altered under conditions of prolonged doubling time. In cells grown in MOPS/acetate, the β -lactamase concentration increases more than threefold, whereas the cellular concentration of CAT falls to two-thirds of its value in L-broth (Table 2). Bacteria grow in MOPS/acetate at one-fifth the rate seen in L-broth; by multiplying these factors, we calculate that the rate of synthesis of β -lactamase must be nearly independent of growth rate, whereas the synthesis of CAT must slow by a factor of ~ 8 in MOPS/acetate. As the steady-state concentration of *bla* mRNA is hardly influenced by growth rate (see above), it is not surprising that the rate of production of β -lactamase is similarly unaffected. In contrast, the substantial decline in the rate of synthesis of CAT in cells growing in MOPS/acetate can be accounted for by the observed decrease in the cellular concentration of *cat* mRNA, which is seen here to result primarily from a decrease in the half life of the *cat* transcripts.

The results reported here indicate that the decay rates of various messages are affected differently by changes in the rate of cell growth. Growth-related effects on decay occur with both a very stable mRNA species (*ompA*) and an mRNA species (*cat*) that has a half life of 2–3 min, as is typical of most *E. coli* mRNAs. The effect of growth on mRNA decay does not appear to be correlated with either the extent of mRNA secondary structure or the nature of the gene product. The *ompA* and *lpp* transcripts both possess the potential for extensive intramolecular basepairing, and both encode abundant outer membrane proteins; yet the decay of only one of these messages (*ompA*) is modulated in response to changes in the bacterial growth rate.

Our observations indicate that control of the rate of mRNA decay is one means by which cells can reduce the rate of biosynthesis of proteins during times of slow growth. Moreover, the finding that only certain *E. coli* messages are degraded more rapidly at slow bacterial growth rates implies the existence of mechanisms that differentially alter the decay rates of transcripts.

We thank Staffan Normark and Ulf Henning for stimulating discussions. These studies were supported by the Riksförningen mot Cancer of Sweden and the US NIH.

Received 2 July; accepted 2 August 1984.

1. Pederson, S., Bloch, P. L., Reeh, S. & Neidhardt, F. C. *Cell* **14**, 179–190 (1978).
2. Sarmientos, P., Sylvester, J. E., Contente, S. & Cashel, M. *Cell* **32**, 1337–1346 (1983).

3. Nomura, M., Yates, J. L., Dean, D. & Post, L. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7084–7088 (1980).
4. Beck, E. & Bremer, E. *Nucleic Acids Res.* **8**, 3011–3024 (1980).
5. Movva, N. R., Nakamura, K. & Inouye, M. *J. molec. Biol.* **143**, 317–328 (1980).
6. Alton, N. K. & Vapnek, D. *Nature* **292**, 864–869 (1979).
7. Nakamura, K. & Inouye, M. *Cell* **18**, 1109–1117 (1979).
8. Sutcliffe, J. G. *Cold Spring Harb. Symp. quant. Biol.* **43**, 77–90 (1978).
9. von Gabain, A., Belasco, J. G., Schottel, J. L., Chang, A. C. Y. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **80**, 653–657 (1983).
10. Lugtenberg, B., Peters, R., Bernheimer, H. & Berendsen, W. *Molec. gen. Genet.* **147**, 251–262 (1976).
11. Ingraham, J. L., Maaloe, O. & Neidhardt, F. C. *Growth of the Bacterial Cell* (Sinauer Associates, Sunderland, Massachusetts, 1983).
12. Apirion, D. *Processing of RNA* (ed. Dunn, J. J.) 286–306 (Brookhaven National Laboratory, Upton, New York, 1974).
13. Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
14. Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134**, 1141–1156 (1978).
15. Stueber, D. & Bujard, H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 167–171 (1981).
16. Seed, B. *Nucleic Acids Res.* **10**, 1799–1810 (1982).
17. Lee, N., Nakamura, K. & Inouye, M. *J. Bact.* **146**, 861–866 (1981).
18. Fisher, J., Belasco, J. G., Khosla, S. & Knowles, J. R. *Biochemistry* **19**, 2895–2901 (1980).
19. Gorman, C. M., Moffat, L. F. & Howard, B. H. *Molec. cell. Biol.* **2**, 1044–1051 (1982).
20. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
21. Stueber, D. & Bujard, H. *EMBO J.* **1**, 1399–1404 (1982).
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
23. Thomas, P. S. *Meth. Enzym.* **100**, 255–266 (1983).
24. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–723 (1977).
25. LeGrice, S. F. J., Matzura, H., Marcoli, R., Iida, S. & Bickle, T. A. *J. Bact.* **150**, 312–318 (1982).

Synthesis in yeast of a functional oxidation-resistant mutant of human α_1 -antitrypsin

Steven Rosenberg, Philip J. Barr, Richard C. Najarian & Robert A. Hallelwell

Chiron Research Laboratories, Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

Cumulative damage to lung tissue by leukocyte elastase is thought to be responsible for the development of pulmonary emphysema, an irreversible lung disease characterized by loss of lung elasticity^{1–3}. It is also thought to be involved in the rapidly developing and usually fatal adult respiratory distress syndrome⁴. The primary defence against elastase damage is the anti-protease known as α_1 -antitrypsin^{5,6}, a glycosylated serum protein of 394 amino acids³. Oxidation of the methionine 358 residue located at the active centre of α_1 -antitrypsin results in a dramatic decrease in inhibitory activity towards elastase which effectively inactivates the protective function⁶. It has been suggested that this oxidation sensitivity has a regulatory function and allows tissue breakdown at sites of inflammation by inactivation of α_1 -antitrypsin by oxygen radicals released by phagocytes³. In the above diseases, however, the oxidative inactivation of α_1 -antitrypsin is probably of major importance in allowing lung damage by elastase^{4,7}. An oxidation-resistant α_1 -antitrypsin might make it possible to reduce the large doses of α_1 -antitrypsin required for emphysemics and provide treatment for acute inflammatory respiratory conditions. To further the possibility of therapy for the above conditions, we describe here the synthesis in yeast of active, non-glycosylated, human α_1 -antitrypsin. Site-directed mutagenesis has been used to construct an active, oxidation-resistant derivative containing a single methionine to valine substitution at the active centre. This demonstrates the potential of engineered modifications to protein molecules designed to improve their physiological function.

To obtain α_1 -antitrypsin cDNA sequences encompassing the entire coding region, a cDNA library constructed from poly(A) mRNA from liver was screened with radiolabelled oligonucleotide probes corresponding to amino acid residues –23 to –17 (25,000 colonies) and 344–350 (5,000 colonies). The probe sequences were taken from the partial nucleotide sequences described previously^{8,9}. Four colonies hybridized^{10,11} to the N-terminal probe and 148 to the C-terminal probe. After gel

Fig. 1 Nucleotide sequence and predicted amino acid sequence of α_1 -antitrypsin cDNA encoding the pre-protein. For expression in yeast the pre-sequence was removed by *Bam*HI digestion. The reactive centre Met-Ser at positions 358–359 is boxed. Amino acids above the sequence identify differences between this protein sequence and those derived from *a*, protein sequencing³; *b*, the cDNA of Woo *et al.* (see ref. 3); *c*, the cDNA of Bollen *et al.*³⁵.

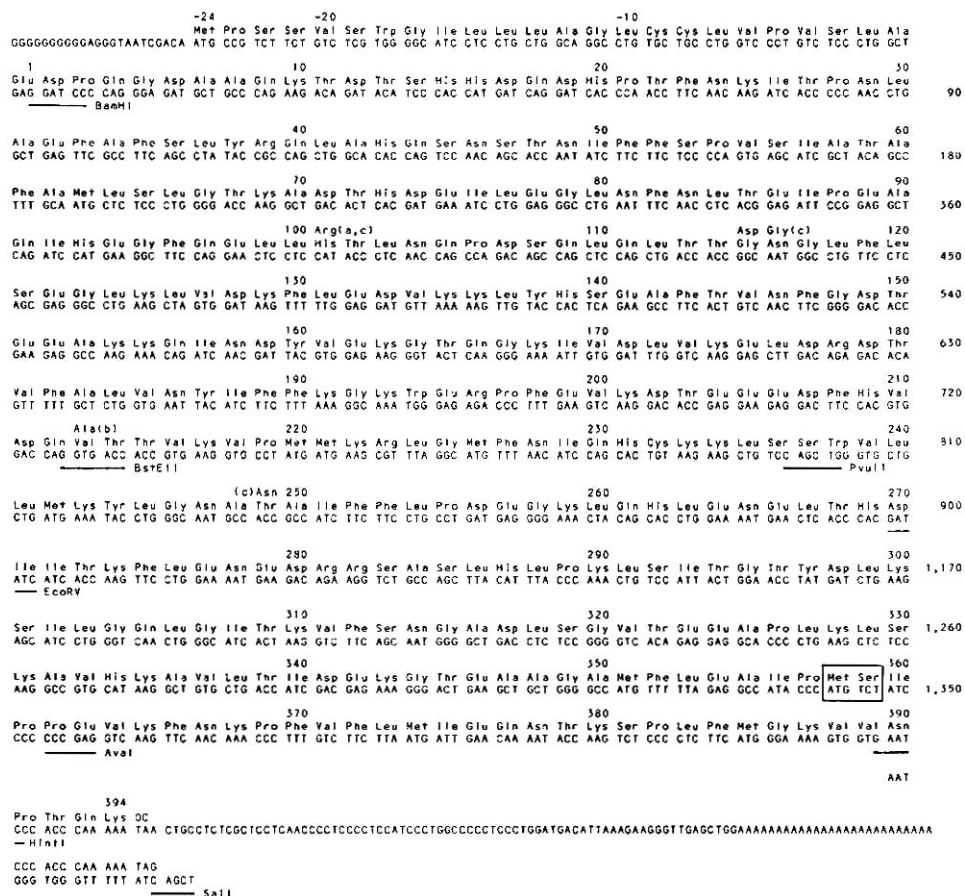
Methods: Double-strand cDNA was synthesized as described previously³⁶, size fractionated on a Sepharose CL4B column to remove most of the cDNA of <300 bp, treated with *S*₁ nuclease and cloned at the *Pst*I site of a pBR322 derivative using the homopolymeric dG and dC tailing method³⁷. Note that the second-strand cDNA was self-primed; the first seven nucleotides of the cDNA differ from the genomic sequence and are probably incorrect as has been described previously³⁸. The following restriction fragments shown were isolated by preparative polyacrylamide gel electrophoresis³⁷: 603 bp *Bam*HI–*Bst*E2 from p5'AT-1; 450 bp *Bst*E2–*Ava*I from p3'AT-11; 85 bp *Ava*I–*Hin*FI from p3'AT-3. The five C-terminal amino acids and termination codon were encoded by the synthetic oligonucleotide adapters shown adjacent to the coding sequence. These oligomers were synthesized by phosphoramidite chemistry²⁸, phosphorylated³⁷, and mixed in equimolar amounts. Approximately 2 pmol of fragments 1 and 2 were ligated³⁷ together in 50 μ l and after removal of ligase, digested with *Bam*HI and *Ava*I. Similarly, fragment 3 and the synthetic duplex were ligated and digested with *Ava*I and *Sal*I. The resulting fragment mixtures were ligated, digested with *Bam*HI and *Sal*I and DNA of ~1,200 bp gel isolated and cloned by substitution into *Bam*HI, and *Sal*I cut and phosphatased³⁷ pBR322. Analysis of miniscreen plasmid DNA³⁷ from these transformants indicated that a full length cDNA had been obtained and one such transformant was used in subsequent constructions.

analysis of plasmid DNA from the 4 N-terminal and 12 C-terminal clones, three overlapping clones were selected for further study. We have elucidated a composite DNA sequence for α_1 -antitrypsin derived from the three cDNA plasmids by subcloning and sequencing in M13 phage vectors^{12,13} (Fig. 1). The protein sequence predicted by the cDNA differs from the published sequences in several positions, probably due to the genetic polymorphisms known to be quite common in α_1 -antitrypsin¹⁴.

To assemble a cDNA encoding the mature α_1 -antitrypsin suitable for expression in yeast, three restriction fragments from the *Bam*HI–*Hin*FI sites shown in Fig. 1 were first isolated from their respective plasmids. These fragments were ligated together with the synthetic DNA segment shown in Fig. 1 and cloned by substitution between the *Bam*HI and *Sal*I sites of pBR322¹⁵.

To construct a yeast plasmid directing the synthesis of α_1 -antitrypsin we first isolated the yeast acid phosphatase (PHO5) promoter from a library of yeast genomic DNA using a synthetic oligonucleotide probe (Fig. 2). The PHO5 promoter was chosen because it is a relatively strong yeast promoter which can be regulated using either phosphate concentration¹⁶ or temperature-inducible mutants¹⁷. The second codon of the acid phosphatase gene was joined in phase to the second codon of the α_1 -antitrypsin cDNA by blunt-ended ligation (Fig. 2). After subcloning, this PHO5– α_1 -antitrypsin segment was isolated as a *Bam*HI–*Sal*I fragment and joined to a *Sal*I–*Bam*HI fragment containing the transcription terminator of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene (ref. 18 and our unpublished results), followed by insertion into the pC1/1 yeast shuttle vector as a *Bam*HI fragment.

To change the methionine residue at position 358 to a valine residue, we used M13 site-directed mutagenesis¹⁹. The *Pvu*II–



*Sal*I fragment of the cDNA (Fig. 1) coding for amino acids 238–394 was cloned by *Sma*I–*Sal*I substitution into M13 mp10¹³. Single-stranded viral DNA was isolated and used as a template for site-directed mutagenesis using a synthetic 30-mer as a mutagenic primer (Fig. 3). Positive plaques were identified using the primer as a probe and the codon change was confirmed by DNA sequencing (Fig. 3). The 360 base pair (bp) *Eco*RV–*Sal*I fragment of α_1 -antitrypsin (Fig. 1) containing the mutation, was inserted into the yeast expression plasmid as described in Fig. 3 legend. This fragment was re-sequenced from pC1/1PHO5AT(Val), confirming the T to C transition and showing that no other mutations had occurred during the M13 mutagenesis, as have been observed previously²⁰. Yeast strain AB103²¹ was transformed²² with plasmids pC1/1, pC1/1PHO5AT(Met) and pC1/1PHO5AT(Val).

To characterize the α_1 -antitrypsin produced in yeast, cultures were grown in high (repressing) or low (inducing) phosphate and collected in log phase. Cells were lysed with glass beads, and the soluble cell proteins analysed by SDS-polyacrylamide gel electrophoresis and elastase inhibition assays (Fig. 4; Table 1). The results of the gel analysis (Fig. 4A) show that a protein of ~42,000 molecular weight (MW) is specifically induced in those transformants containing the α_1 -antitrypsin cDNA. A protein of 44,000 MW is expected for mature, non-glycosylated α_1 -antitrypsin. Although strain AB103 contains the *pep*4-3 mutation and is deficient in most of the proteolytic activities of the yeast vacuole²³, additional smaller proteins of ~23,000 and 19,000 MW were specifically induced. These were shown to be antigenically related to α_1 -antitrypsin by Western blot analysis (Fig. 4B) and possibly result from proteolytic degradation. When synthesis of α_1 -antitrypsin was similarly analysed in strain

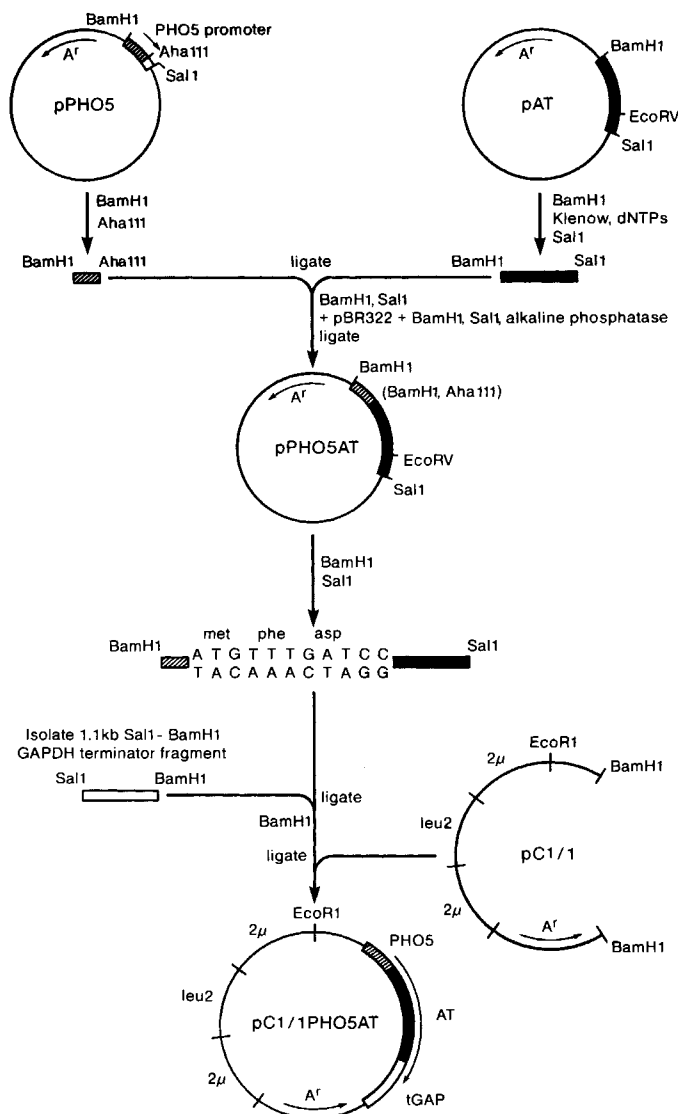


Fig. 2 Construction of yeast plasmids expressing a PHO5- α_1 -antitrypsin fusion encoding the first two amino acids of the PHO5 gene and residues 2-394 of α_1 -antitrypsin. DNA fragments and plasmids are only approximately to scale. To isolate the PHO5 promoter, the oligonucleotide probe $5'$ GGCACT-CACACGTGGGACTAG $3'$, complementary to a portion of the sequence of the promoter, was used to screen a library of *Saccharomyces cerevisiae* DNA³⁹. Several positive clones were obtained and one was sequenced^{12,13} and shown to contain the PHO5 promoter. A 630 bp *Bam*HI-*Sal*I fragment from this isolate was subcloned into pBR322. This fragment contains 80 bp of the acid phosphatase coding sequence and has the sequence $5'$ ATGTTTAA $3'$ encoding the first three amino acids, the second and third codon specifying an *Aha*III site. pC1/1 is a derivative of pJDB219 (ref. 40) in which the pMB9 sequences have been replaced by pBR322.

2150 *pep4*⁺, an increased level of the 23,000 and 19,000 MW fragments was observed as well as the production of a new fragment of 40,000 MW (data not shown).

To determine the activities of α_1 -antitrypsin(Met) and α_1 -antitrypsin(Val), appropriate yeast extracts were incubated with human leukocyte elastase and a synthetic peptide substrate whose cleavage can be assayed by increase in absorbance at 410 nm⁶. Inhibition of elastase was found to be directly proportional to the amount of human α_1 -antitrypsin added in the presence or absence of control yeast extracts (Table 1 assays 1-5, 13, and our unpublished results). Control yeast extracts had no effect on elastase activity (Table 1, assay 6) whereas

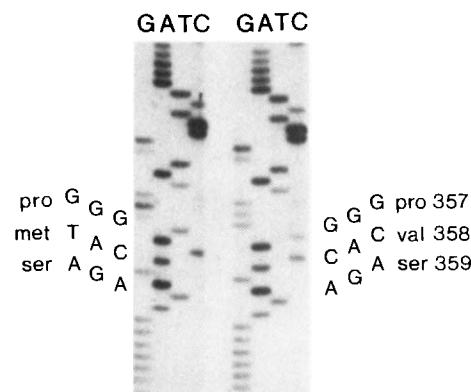


Fig. 3 Autoradiogram of the DNA sequencing gel showing the nucleotide sequence of wild-type (Met) and mutant (Val) active centres of α_1 -antitrypsin.

Methods: A mutagenic primer having the sequence $5'$ GACCTCGGGGGGATAGACA $3'$ was synthesized and used to prime second-strand DNA synthesis on single-stranded M13 containing a wild-type α_1 -antitrypsin fragment (see text) as described previously^{19,21}. The mismatch is underlined. Mutagenized phage were located using the same primer as a probe and filters were washed for 60 min at 65 °C in 0.15 M NaCl, 15 mM sodium citrate (1×SSC). To substitute the Met to Val mutant in the yeast expression vector shown in Fig. 2, the 360 bp *Eco*RV-*Sal*I fragment of α_1 -antitrypsin (Fig. 1), containing the mutation, was prepared from a digest of the M13 replicative form DNA. This was ligated to the 1,350 bp *Bam*HI-*Eco*RV fragment, containing the PHO5 promoter and N-terminal residues of α_1 -antitrypsin, prepared from pPHO5AT, and the 1,100 bp *Sal*I-*Bam*HI GAPDH terminator fragment, both shown in Fig. 2. After ligation, the mixture was digested with *Bam*HI before ligation into *Bam*HI-cut and phosphatased pC1/1. After transformation of HB101⁴¹, one positive clone, analogous to pPHO5AT(Met), was obtained after restriction enzyme analysis of plasmid DNA from 12 clones.

Table 1 Effects of *N*-chlorosuccinimide oxidation on α_1 -antitrypsin activity in yeast extracts assayed by inhibition of elastase activity

Assay no.	Plasmid	Volume of yeast extract (μl)	μg α_1 -Antitrypsin	NCS (μl) (10 mM)	Elastase activity (%)
1	—	—	—	—	100
2	—	—	0.1	—	85
3	—	—	0.2	—	67
4	—	—	0.5	—	22
5	—	—	1.0	—	0
6	C1/1	1.5	—	—	100
7	Met	0.5	—	—	54
8	Met	1.5	—	—	8
9	Val	0.5	—	—	52
10	Val	1.5	—	—	15
11	—	—	1.0	—	5
12	—	—	1.0	10	91
13	C1/1	3.8	1.0	—	12
14	C1/1	3.8	1.0	10	84
15	Met	1.5	—	—	8
16	Met	1.5	—	10	93
17	Val	1.5	—	—	20
18	Val	1.5	—	10	18

Yeast extracts were prepared by lysis with glass beads and centrifuged to remove cellular debris³². Extracts were diluted in 50 mM Tris pH 8, 50 mM NaCl. For *N*-chlorosuccinimide (NCS) oxidations, 7.5 μl of extracts were incubated at room temperature for 5 min in 100 μl of 50 mM Tris pH 8, with the indicated amount of NCS (10 mM). Aliquots were transferred to tubes containing 0.1 mM Meo-Suc-Ala-Ala-Pro-Val-*p*-nitroanilide (Vega), 50 mM Tris pH 8, 0.5 M NaCl, 100 μg bovine serum albumin (BSA), in a volume of 1 ml. Assays were initiated by the addition of 0.15 μg HLE and incubated for 15 min at 28 °C, terminated by the addition of 100 μl of 8 M acetic acid and the absorbance at 410 nm was determined⁶. Protein concentrations³³ were: C1/1, 6.4 mg ml⁻¹; Met, 8.0 mg ml⁻¹; Val, 9.7 mg ml⁻¹. From the amount of human leukocyte elastase (HLE) activity observed, given the protein concentration of the yeast extracts and the amount of HLE added, the percentage of the yeast-soluble protein that is α_1 -antitrypsin can be calculated. We assume 100 % HLE activity, a 1:1 complex between HLE and α_1 -antitrypsin, and that the molecular weights of HLE³⁴ and α_1 -antitrypsin are 29,000 and 42,000 MW, respectively. In this experiment α_1 -antitrypsin is 3.0 and 3.5 % in the Met extract (assays 7 and 8) and 2.3 and 2.6 % in the Val extract (assays 9 and 10).

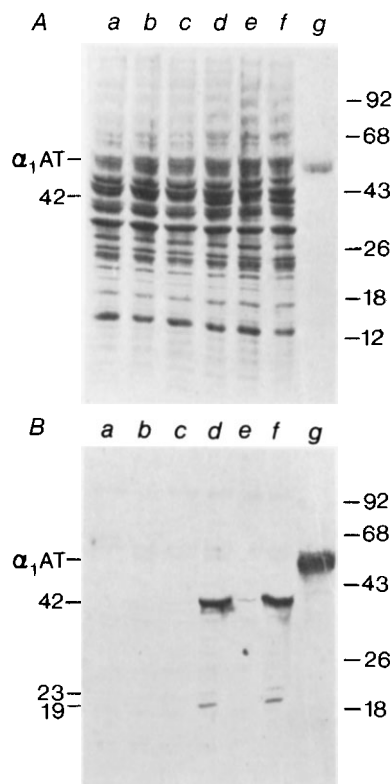


Fig. 4 α_1 -antitrypsin (α_1 AT) and related proteins synthesized in yeast. **A**, SDS 10–20% polyacrylamide gradient gel⁴² stained with Coomassie blue showing soluble proteins synthesized in: *a*, uninduced pC1/1-containing control cells; *b*, induced pC1/1-containing control cells; *c*, uninduced pC1/1PHO5(Met) cells; *d*, induced pC1/1PHO5(Met) cells; *e*, uninduced pC1/1PHO5(Val) cells; *f*, induced pC1/1PHO5(Val) cells; *g*, glycosylated α_1 -antitrypsin from human serum. **B**, Proteins from gel **A** which react specifically with antibody to α_1 -antitrypsin. Identical extracts were electrophoresed on the same gel, transferred to nitrocellulose paper and α_1 -antitrypsin-related proteins visualized by antibody-conjugated horseradish peroxidase staining⁴³. Elastase inhibition assays (Table 1) indicated that induced extracts contained 1.0 % α_1 -antitrypsin in this experiment.

Methods: Yeast cells were grown in 50 ml cultures of leu⁻ selective media in high or low phosphate to an A_{650} of ~2. After lysis³², ~20 μ g of the soluble proteins were electrophoresed on a 10–20 % SDS-polyacrylamide gradient gel and blotted onto nitrocellulose paper. α_1 -Antitrypsin-related proteins were visualized by washing the nitrocellulose paper with rabbit anti- α_1 -antitrypsin (Accurate Chemicals) followed by horseradish peroxidase/goat anti-rabbit antibody staining⁴³. Media containing high and low levels of phosphate consisted of synthetic complete medium lacking leucine⁴⁴ supplemented with 0.2 volumes of phosphate-depleted YEP medium⁴⁵, 2 g l⁻¹ asparagine replacing (NH₄)₂SO₄, 1 g l⁻¹ KH₂PO₄ (high phosphate) or 1 g l⁻¹ KCl, 0.03 g l⁻¹ KH₂PO₄ (low phosphate). Yeast strains are AB103.1 (*Mata leu2-3, 112 ura3-52 pep4-3 his4-580 [cir^o]*) and 2150-2-3 (*MATa leu2 ade1 [cir^o]*).

induced yeast extracts for α_1 -antitrypsin(Met) and α_1 -antitrypsin(Val) both inhibited elastase activity (Table 1, assays 7–10). The percentage of α_1 -antitrypsin in yeast extracts (Table 1) showed approximate agreement with estimates from SDS-polyacrylamide gels, indicating normal activity of the recombinant α_1 -antitrypsins.

To discover if α_1 -antitrypsin(Val) was oxidation resistant, we treated extracts with *N*-chlorosuccinimide, which converts susceptible methionines to their sulfoxides²⁴. Both the recombinant α_1 -antitrypsin(Met) and a reconstituted control extract containing human α_1 -antitrypsin are susceptible to oxidative inactivation whereas the α_1 -antitrypsin(Val) is resistant (Table 1, assays 13–18).

It has been reported recently that the development of emphysema can be limited by intravenous injection of α_1 -antitrypsin partially purified from human serum²⁵. Although this enzyme is the most abundant anti-protease in human serum (~1.3 g l⁻¹)²⁶, insufficient amounts are available from plasma to treat all affected individuals because its half life is only 6 days²⁷. Production from yeast cells offers a viable alternative, especially as yields of >10 % of soluble cell protein as α_1 -antitrypsin have been obtained recently using the glyceraldehyde phosphate dehydrogenase promoter²⁸ (manuscript in preparation). It will, however, be important to establish whether the half lives in serum and the rates of inhibition are comparable for the non-glycosylated recombinant α_1 -antitrypsins. The replacement of methionine by valine at the active centre of the molecule was based on the observation that a synthetic peptide with this substitution is an improved substrate for elastase²⁹. Recent studies have shown that a naturally occurring arginine substitution at this position lacks elastase inhibitory activity but has gained the specificity of antithrombin III³⁰. Thus, it is important to determine the specificities of any derivatives with therapeutic potential. In addition to providing improved treatment for pathogenic deficiencies of α_1 -antitrypsin, the construction of novel α_1 -antitrypsin derivatives by protein engineering, together with kinetic⁶ and crystallographic³¹ studies, should provide a better understanding of the regulation and reaction mechanism of this important anti-protease.

We thank James Travis for gifts of α_1 -antitrypsin and HLE, James Travis and Robin Carrell for useful discussions, Robert Huber for ref. 31, Qui Lim Choo, Kathy Mead, Bernardita Mendez, Patricia Tekamp-Olson, and other colleagues at Chiron for help with experiments and the manuscript.

Received 10 May; accepted 30 August 1984.

1. Eriksson, S. *Acta med. scand.* **175**, 197–205 (1964).
2. *Lancet* **i**, 743–745 (1980).
3. Carrell, R. W. *et al. Nature* **298**, 329–334 (1982).
4. Cochrane, C. G., Spragg, R. G., Revak, S. D., Cohen, A. B. & McGuire, W. W. *Am. Rev. resp. Dis.* **127**, Suppl. 2, 25–27 (1983).
5. Laurell, C.-B. & Eriksson, S. *Scand. J. clin. Lab. Invest.* **15**, 132–140 (1963).
6. Beatty, K., Bieth, J. & Travis, J. J. *biol. Chem.* **255**, 3931–3934 (1980).
7. Gadek, J. E., Fells, G. A. & Crystal, R. G. *Science* **206**, 1315–1316 (1979).
8. Kurachi, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6826–6830 (1981).
9. Leicht, M. *et al. Nature* **297**, 655–659 (1982).
10. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6613–6617 (1981).
11. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989–2998 (1981).
12. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
13. Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
14. Fagerhol, M. K. & Cox, D. W. *Adv. hum. Genet.* **11**, 1–62 (1981).
15. Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
16. Meyhack, B., Bajwa, W., Rudolph, H. & Hinnen, A. *EMBO J.* **6**, 675–680 (1982).
17. Kramer, R. A., DeChiara, T. M., Schaber, M. D. & Hilliker, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 367–370 (1984).
18. Holland, M. J. & Holland, J. P. *J. biol. Chem.* **254**, 5466–5474 (1979).
19. Zoller, M. J. & Smith, M. *Nucleic Acids Res.* **10**, 6487–6500 (1982).
20. Villafranca, J. E. *et al. Science* **222**, 782–788 (1984).
21. Brake, A. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4642–4646 (1984).
22. Hinnen, A., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929–1933 (1978).
23. Hemmings, B. A., Zubenko, G. S., Hasilik, A. & Jones, E. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 435–439 (1981).
24. Johnson, D. & Travis, J. J. *biol. Chem.* **253**, 7142–7144 (1978).
25. Gadek, J. E. & Crystal, R. G. *Am. Rev. resp. Dis.* **127**, Suppl. 2, 45–46 (1983).
26. Jeppsson, J.-O., Laurell, C.-B. & Fagerhol, M. K. *Eur. J. Biochem.* **83**, 143–153 (1978).
27. Jeppsson, J.-O., Laurell, C.-B., Nossin, B. & Cox, D. W. *Clin. Sci. molec. Med.* **5**, 103–107 (1978).
28. Urdea, M. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 7461–7465 (1983).
29. Nakajima, K., Powers, J. C., Ashe, B. M. & Zimmerman, M. J. *biol. Chem.* **254**, 4027–4032 (1979).
30. Owen, M. C., Brennan, S. O., Lewis, J. H. & Carrell, R. W. *New Engl. J. Med.* **694**–698 (1983).
31. Loebermann, H., Tokuko, R., Deisenhofer, J. & Huber, R. *J. molec. Biol.* (in press).
32. Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. & Hall, B. D. *Nature* **298**, 347–350 (1982).
33. Bradford, M. *Analyt. Biochem.* **72**, 248–253 (1976).
34. Baugh, R. J. & Travis, J. *Biochemistry* **15**, 836–841 (1976).
35. Bollen, A. *et al. DNA* **2**, 255–264 (1983).
36. Martial, J. A., Hallewell, R. A., Baxter, J. D. & Goodman, H. M. *Science* **205**, 602–607 (1979).
37. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
38. Porter, A. G. *et al. Nature* **282**, 471–477 (1979).
39. Burke, R. L., Tekamp-Olson, P. & Najarian, R. J. *biol. Chem.* **258**, 2193–2201 (1983).
40. Beggs, J. D. *Nature* **275**, 104–109 (1978).
41. Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41**, 459–472 (1969).
42. Laemmli, U. K. *Nature* **227**, 680–685 (1977).
43. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4355 (1979).
44. Sherman, F. *et al. Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, New York, 1982).
45. Rubin, G. M. *Meth. Cell Biol.* **12**, 45–64 (1975).

Giving 'em the business

W.F. Clocksin

The AI Business: Commercial Uses of Artificial Intelligence.

Edited by Patrick H. Winston and Karen A. Prendergast.

MIT Press: 1984. Pp.324. \$15.95, £14.25.

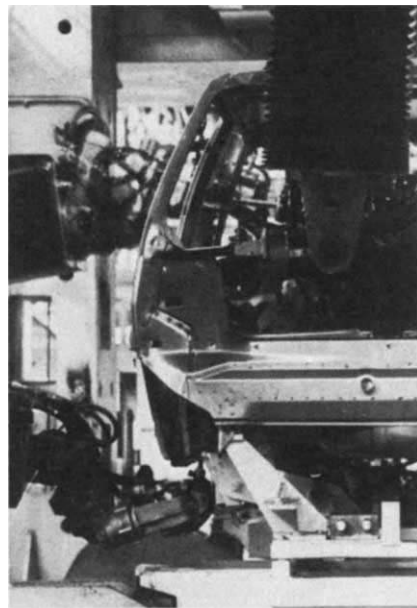
THE past, it seems, is not always another country. In the 1960s, over-optimistic advocacy by researchers of the nascent field of artificial intelligence (AI) created expectations impossible to satisfy. In the 1980s, spurred by the prospect of serious international competition, we again have enthusiastic promotion of the technology of AI — especially expert systems and “fifth-generation” computers — and it is no less fervent: but this time the promoters move in commercial circles. The passionate propaganda from the marketing men dismays the researcher and perplexes the technical manager; as in the 1960s, the hyperbole obscures a realistic assessment of the true potential of AI technology.

In an attempt to clear the air, a colloquium was called by the MIT Industrial Liaison Program to bring together representatives of the four main groups involved: the academics, who started it all; the investment bankers and venture capitalists; those concerned with industrial research and development; and those in industry who use techniques devised by AI researchers because there is a job to be done. *The AI Business* is a collection of the edited transcripts of the colloquium. The book opens and closes with perspective and summary chapters by the editors. The first chapter sets the scene by placing MIT at the centre of AI research, and by giving a short “history” of the field complete with its own Prehistory, Dawn Age, Renaissance and so forth. The practising researcher may find it difficult to take this chapter seriously, while others may be dangerously misled by such cavalier treatment. Similarly, while the last chapter does a fine job of summing up, and offers some homespun advice on what action should be taken by universities and companies, there is again the tendency to lapse into dottiness, viz., “...strength in Artificial Intelligence ensures strength at everything that Artificial Intelligence can help with. And that may be just about everything”. These two chapters are very much the bread of the sandwich though; the meat is proved by the 16 other contributors.

Although Winston (and later, Minsky) admits to regarding the term “expert system” with a certain amount of scorn, the casual reader will not understand why, for the first section of the book contains six chapters describing some of the best examples of work in the field. Here the practitioner will find accounts of his old friends XCON and DIPMETER

ADVISOR, true products of AI research which have reached commercial credibility not by hype, but by quietly saving their users substantial amounts of money.

The next section, entitled “Work and Play”, contains six chapters on a subject close to the hearts of AI researchers: interactive workstations. Part of AI



Robots on the line — automatic welding on the Robogate system at Fiat's Rivalta plant in Italy.

research is about making computers easier to use, particularly as automated design assistants and as apprentices. This goes much farther than the computer-aided draughting found in industry today. The computer should be a participant in the design process, synthesizing candidate designs that conform to sets of constraints with which the designer experiments. There is a fascinating contribution by Alan Kay, who rambles through a compelling mixture of reminiscence and advice. Other chapters cover commercially available systems for electronic design automation and English-language communications with databases. The overall theme of this section is the life-cycle from basic research to exploitation in the marketplace.

The next section deals with robotics. Mike Brady gives a good survey of the relationship of AI research areas to their application in industrial robotics, while the other two contributions survey real-life industrial applications, giving examples of sensor-controlled robots and more general

issues in factory automation. This is the most straightforward and informative section of the book, perhaps as a result of the authors' application of well-understood techniques to well-defined problems, with clear criteria for the success or failure of an application. Other areas of AI are not yet as much under control.

The final section, “Today and Tomorrow”, centres on trends, both technological and economic. Minsky's chapter is a prolonged snort of disapproval at contemporary events in research, politics, business and industry. His basic message, though, is of prime importance. The commercial successes of today rely on AI research carried out ten years ago. Because there is practically no basic work in AI now being done, the future may already be lost. This theme is repeated in the other two contributions by investment analysts. Particularly impressive is the cogent and analytical chapter by the economist, William Janeway. Anyone closely involved with recent Government initiatives in “information technology” in any country will find considerable sympathy with Janeway's view that,

The packaging of ignorance by artifice to obtain finance distorts not only the returns actually received by investors. Distortion of research activity in order to obtain finance on such terms is both a professional and a social cost that we in this country seem ill equipped to appreciate.

The main question to be asked of this book is, “Who is it for?”. As the editors point out in the last chapter, the agreements and differences of opinion expressed at the colloquium have been preserved in the book. The transcripts of the informal discussion sessions at the end of each section are particularly effective in this connection. Great stuff for the practising researcher, who will be intimately familiar with the elusive issues involved, and who will thus be in a position to read between the lines. However, the discussions, together with the editors' own contributions, give the book a tone of superficiality and off-handedness which is dangerous and inappropriate for a more general audience. In spite of this, some of the individual chapters are particularly sound, and deserve wider exposure.

I want to see technical managers and engineers becoming acquainted with the issues covered in this book. But I would tell them not to believe everything they read in it. The simplistic, shoot-from-the-hip flavour may be instinctively appealing, but it endangers credibility, and should not be interpreted as solid worldly-wise experience. Despite Winston's “Six Ages of Artificial Intelligence”, 20 years is not a very long time. What we need, above all, is a little patience. □

W. F. Clocksin is an Assistant Director of Research at the Computer Laboratory, University of Cambridge.

Ways with MCAs

Karol Sikora

Monoclonal Antibodies: Probes for the Study of Autoimmunity and Immunodeficiency.

Edited by Barton F. Haynes and George S. Eisenbarth.

Academic: 1984. Pp.318. \$45, £33.

Monoclonal Antibodies: Principles and Practice.

By James W. Goding.

Academic: 1984. Pp.276. \$22.50, £17.

THE immune system has evolved into one of nature's most perfect defences. Recognition, memory, specificity and a plethora of effector mechanisms all combine to create a powerful and dynamic defence network. But things can go wrong. In autoimmune disease the immune response turns its attention to host tissues and organs causing self-destruction. In immunodeficient states, whether congenital or acquired, the host becomes susceptible to a wide range of normally innocuous infectious agents as well as to an increased risk of malignancy. The problem in dissecting out the immune system's game plan is the multiplicity of players. Monoclonal antibodies (MCAs) have changed this. These highly specific molecular flags enable the precise identification of the system's components; the lymphocyte.

For the first time in one book, Barton Haynes and George Eisenbarth have brought together a collection of reviews on the use of MCAs in the analysis of immunological disorders. Beginning with a summary of the study of T-cell differentiation using the well-known anti-T-lymphocyte subset MCAs, the book proceeds through a series of accounts of haemopoiesis, thymic differentiation, human hybridoma formation and immunodeficiency. The final four chapters discuss classic endocrine diseases such as diabetes, thyroid problems and myasthenia gravis.

The net result is a rather hotchpotch collection. To the immunologist there is little new that has not been reviewed elsewhere, often by the same authors. The clinician searching for a clearing in the increasingly complex immunological jungle will also be disappointed; haematologists, rheumatologists and endocrinologists will not find much of relevance here to their current practices

More on monoclonals

- *Monoclonal Antibodies to Receptors: Probes for Receptor Structure and Function*, edited by M.F. Greaves. Publisher is Chapman & Hall, price is £35, \$75.

- *Monoclonal Antibodies and Functional Cell Lines: Progress and Applications*, edited by R.H. Kennett, K.B. Bechtol and T.J. McKearn, a sequel to an earlier volume of 1982. Publisher is Plenum, price is \$59.40, £47.03.

and how to improve them. Too much data is presented in many chapters instead of the critical reviews which would have been far more helpful. An exception is the chapter on diabetes, which succinctly covers the role of the immune system in the causation of the disease. The editors might consider using this contribution as a model for an improved future edition.

I can be a great deal more enthusiastic about the second book reviewed here, that by James Goding. Although MCAs were discovered nearly a decade ago, there are few practical guides to the experimenter setting up a monoclonal factory. Goding is a scientist of considerable experience in the MCA business, and has written such a guide, a good one.

After a chapter on the theory behind MCA production, the remainder of the book goes step by step through the various stages of production, purification, radiolabelling, the application of MCAs in assays and in immunohistology also being discussed. Particularly pleasing is the step-wise layout of the various practical

procedures described, together with the extensive reference material provided. Of course, no two workers use precisely the same approach as outlined but for the newcomer to the field the recipes given do provide a starting point.

Surprisingly little information, however, is given to guide the neophyte in choosing which myeloma system to use, often a major stumbling block. Several are mentioned but no comparison between them is given. And there is no discussion of the more recently developed forms of immunization using peptide fragments; by preparing MCAs to peptides synthesized on data predicted from DNA analysis, the circle between gene and protein can now be closed.

Altogether, though, this is an excellent laboratory manual, invaluable to the newcomer to the field. But new editions will be required frequently if it is to remain useful over the next few years. □

Karol Sikora is Director of the Ludwig Institute for Cancer Research, Cambridge, UK.

Climate of quality

J.M. Walker

Climates of the Oceans. World Survey of Climatology, Vol. 15.

Edited by H. van Loon.

Elsevier: 1984. Pp. 716. \$173, Dfl. 450.

FEW books can have had as difficult a gestation period as *Climates of the Oceans*. The first editor died, the second withdrew, and when Harry van Loon took over, in 1974, 11 years after the book was commissioned, only one chapter was available in a form suitable for publication. New authors had to be found for four of the other six chapters and, furthermore, Dr van Loon had to accept the constraints imposed by the editorial decisions of his predecessors. The publishers must have wondered if the book was jinxed. Fortunately, they persevered.

The result is an academic textbook and work of reference worthy of its place in the *World Survey of Climatology* series. Moreover, the book has admirably filled a gap in the literature; until now there has been no authoritative and comprehensive survey of weather and climate over the oceans. For that reason it will find a wide audience, including not only meteorologists, oceanographers and geographers, but also those who need information about marine weather and climate for commercial purposes.

As is almost inevitable in a book of this kind, the chapters vary in length, quality and style, but the overall standard of writing and presentation is high. Paradoxically, given the paucity of observational data from most oceanic regions south of the Equator, the longest

chapters are those dealing with the oceans of the Southern Hemisphere. O Höflich's contribution on the climate of the South Atlantic occupies 191 pages; N.A. Streten and J.W. Zillman's (South Pacific) 167; and J.J. Taljaard and H. van Loon's (Indian Ocean south of 35°S) 97. In contrast, M.A. Einarsson's chapter on the climate of Iceland takes up only 25 pages (and the reader is left to guess why it is included in a book on ocean climates!).

It is regrettable that the book is not as up to date as it might have been. It contains few references to works published since 1980. Indeed in "Climate of the Indian Ocean North of 35°S" by C.S. Ramage, the most recent reference is dated 1975, and in K. Terada and M. Hanzawa's chapter on the North Pacific it is 1974. The obvious conclusion is that contributions were submitted over a period of years and were not subsequently updated. This is a pity, because in the past decade there have been numerous studies of atmospheric behaviour over the oceans, particularly over the North Atlantic, the North Pacific and monsoonal parts of the Indian Ocean. In Ramage's article, for example, there is no mention of the Monsoon Experiment (MONEX), while G.B. Tucker and R.G. Barry (who deal with the North Atlantic) say very little about the GARP Atlantic Tropical Experiment (GATE).

All things considered, however, it is remarkable that the shortcomings of the book are so few. It is a work of high quality and a mine of information and data. For anyone with an interest in the oceans it will be invaluable. □

J.M. Walker is Senior Lecturer in Meteorology and Physical Oceanography at the University of Wales Institute of Science and Technology, Cardiff.

Husbanding resources

Colin P. Groves

Evolution of Domesticated Animals.

Edited by I.L. Mason.

Longman: 1984. Pp.452. £35.

As archaeozoology has evolved into a discipline in its own right, so there has been a surge of interest in the economics of domestic animals and in the possibility of domesticating new species. Until recently there was little in the way of synthesis on the subject, and no general books had appeared since Zeuner's heavily archaeologically orientated *A History of Domesticated Animals* (Hutchinson, 1963). Now, suddenly, there is a rush to fill the vacuum. Two recent books of note have been Juliet Clutton-Brock's *Domesticated Animals from Early Times* (Heinemann, 1981), a general book suitable as an undergraduate text, and Helmut Hemmer's *Domestikation: Verarmung der Merkwelt* (Vieweg, 1983), so far not translated into English but an essential read for anyone interested in the background to and consequences of the domestication process itself.

Now comes this compilation, intended, as the introduction says, as

a single-volume account of the origin and history of domestic animals, species by species, which will not only be a ready source of exact information but will also provide an introduction to the literature.

In other words, it is not a theory-orientated book like Hemmer's, nor a textbook like Clutton-Brock's, but a source book of data and literature.

There are 69 chapters on different species or groups, including some on potential domesticates and some on experiments in progress. The longer chapters are divided into standard headings: introduction (with sections on nomenclature, general biology, and numbers and distribution); wild ancestor; domestication and early history; recent history and present status; and future prospects. The treatment of these various themes is very variable, and reflects the main interests of the authors of the chapters. Bökönyi is an archaeologist; nearly half of his chapter on the horse is taken up by the section "Domestication and Early History". Epstein and Bichard are both specialists in animal husbandry; more than half of their chapter on the pig is "Recent History and Present Status". Many other contributions are similarly weighted; perhaps the most successful and rounded of them are those by co-authors with different specialities (such as the chapter by Novoa, a veterinarian, and Wheeler, an archaeologist, on llama and alpaca), or by single authors who have made one or a group of species their own (the chapter on the dog, by Clutton-Brock;

on the reindeer, by Skjenneberg; and the series on domestic fowl, turkey and goose, by Crawford). There is always at least a nod towards all the statutory topics, though in some instances an author may be way out of his depth: Ryder flounders desperately amid the complexities of the taxonomy of wild sheep, and is not always happy with archaeological concepts (it is surprising to see his apparent unquestioning acceptance of high proportions of juveniles as evidence for domestication in an archaeological context, a line of argument which one would have thought long since thrown into disrepute), but as soon as he starts writing about wool, his lifelong speciality, he is clear, authoritative and a joy to read.

The section on general biology is, one would have thought, pivotal to each chapter, but perhaps because it is part of the introduction it is in most instances disappointingly skimmed over. Some of the best accounts of species' biology are, unexpectedly, in the chapters on minor species (Bali cattle, by Rollinson; yak, by Bonnamaire).

My other main criticism of the book is its lack of an overview. Charles Reed does provide an introductory chapter, "The Beginnings of Animal Domestication", but this is of course from an archaeological point of view, and one would have liked to see similar chapters to cover themes such as the biology of domestication; the question of parallel changes in different species as a consequence of domestication; the suitability of different species for domestication; prospects for new domestication; and so on. I may have misunderstood the purpose of the book here, but it seems to me that discussion of such generalities would have been helpful from both a theoretical and a practical standpoint.

When all is said and done, however, it is a pleasure to read experts expounding the mysteries of their art and the book remains a monumental success. It is a compilation of essential data — the only one there is. It deals with all species that have ever been domesticated, or conceivably ever could be: the African elephant is there, the European bison, the onager (though I personally think this ought not to have been included), the ostrich, the crocodile. We have the latest on the cane-rat domestication project from the University of Ibadan group; hamsters and gerbils; snails and oysters; and a fascinating survey of silkworms. Whatever its defects, the book will stand for years to come as an invaluable reference work. It will deservedly find a place in the libraries of universities, animal husbandry and veterinary institutes, archaeology teams, and even many schools. And thank you, Mr Mason, for doing it; because it needed to be done. □

Colin P. Groves is a Senior Lecturer in the Department of Prehistory and Anthropology, Australian National University, Canberra.

ADVERTISEMENT

Announcing a new editor
for an important journal

MODERN GEOLOGY

Editor: L. Beverly Halstead,
University of Reading, UK

MODERN GEOLOGY reports experimental and theoretical findings at the frontier of geological research. Original papers, reviews and short communications afford comprehensive coverage of important developments in the earth and planetary sciences. The editor aims to preserve a balanced perspective of the physical, chemical and biological spheres and to identify major trends and areas of growth in the field. To complete its far-ranging coverage, the journal offers discussion and comments on previously published contributions, as well as book reviews. Another attractive feature of the journal is that it includes pertinent colour plates, published at no charge to the contributing author.

Current & forthcoming papers:

Hot Dry Rock Geothermal Exploitation in the United Kingdom, A. S. Batchelor. Facies Evolution in the Gulf of Suez Area: Sedimentation History as an Indicator of Rift Initiation and Development, B. W. Sellwood & R. E. Netherwood. Early Cretaceous Insects from Southern England, E. A. Jarzembowski. Devonian and Carboniferous Tetrapods — A Global Overview, M. J. Boyd. Molecular Organic Geochemistry — A Tool for the Geosciences, G. Eglinton.

Subscription information:

Current volume block: Volumes 8-9
Price per volume of 4 issues: US \$224.00
50% discount to personal subscribers
ISSN: 0026-7775
FREE SAMPLE COPY AVAILABLE

To: Gordon + Breach, c/o STBS,
42 William IV Street,
London WC2N 4DE, UK
Or: Gordon + Breach, c/o STBS,
50 West 23 Street, New York, NY
10010, USA

Yes, please send me a free sample copy
of **MODERN GEOLOGY**

Name

Affiliation

Address

N

Pollination renewed

Peter D. Moore

Pollination Biology. Edited by Leslie Real. Academic: 1984. Pp.338. Hbk \$49, £36; pbk \$19.50, £15.

"FIGS, yuccas and specialized orchids have proved to be the exception rather than the rule." Thus Peter Raven introduces this compendium of review papers in pollination biology as he seeks to emphasize that it is a subject in which great changes are taking place. The period of descriptive work, from which the extraordinary stories of host-specific fig wasps and yucca moths emerged, is now past and experimental research on population variation and the selective advantage gained by specific individuals with particular attributes has begun.

As in most areas of ecology, the direction in which pollination biology is heading is towards a search for generalities rather than specifics. Can one, for example, make general statements about the evolution of breeding systems in plants in relation to pollination mechanisms? Can one quantify the degree of competition between plants for pollinators? Or is it possible to assess the adaptive significance of a specific trait, such as flower colour? These are the sort of problems which the book addresses.

The approach is radical, in places almost iconoclastic. Many old and well-engrained dogmas are critically examined and discarded, though they are not always replaced with simple alternatives. Take the evolution of dioecy, for example. It has long been assumed that the separation of sexes on to different plants provided advantages in terms of out-breeding. Here Wyatt claims that this principle is not supported by experimental evidence, but what does account for the development of dioecy is still somewhat obscure. Perhaps it is a matter of resource allocation, sexual selection or predation avoidance.

A chapter by Stephenson and Bertin is devoted to sexual selection in plants and approaches the subject in a logical way by dividing it into two categories: (i) competition between members of one sex for reproductive access to the other sex and (ii) choice of a partner by one of the sexes. Too often these groups have been confused in botanical studies; it is not difficult to understand how males may compete for access to females, but less easy to visualize the element of female choice. The latter is perhaps best regarded as non-random fertilization following pollen deposition, that is, some form of incompatibility. But this is questionably comparable to mate selection in animals. There is also the possibility that the female can favour specific male plants or pollinators by such characteristics as a restriction of flowering time. An interesting outcome of these processes is that sexual selection in plants

has not led to the visually obvious secondary sexual characters found in animals, but may often have influenced non-visual, physiological and biochemical developments.

This is a book which tackles many topics of interest to geneticists, ecologists and palynologists. Under what circumstances is wind pollination advantageous? What sensory stimuli determine the foraging behaviour of a pollinating insect? How does the density of a population of flowering plants affect the frequency of visits from a pollinator? These are exciting questions which cut across the boundaries of traditional disciplines.

Viewed as an undergraduate text, which it claims to be in part, the book is not

entirely satisfactory. The text is laden with multiple citations; too many subjects are discussed from different angles in separate chapters; and summaries of chapters are often inadequate as a guide to chapter content. But as a source of information and, above all, stimulating ideas for specialists and research workers, this is an extremely valuable addition to a literature which is currently dominated by detail and description rather than a conceptual approach. Figs and yuccas may be the exceptions, but the mundane may yet prove to be even more enthralling. □

Peter D. Moore is Reader in Ecology in the Department of Plant Sciences, King's College, University of London.

Wholesome genetics

Nick Proudfoot

Eukaryotic Genes: Their Structure, Activity and Regulation.

Edited by N. Maclean, S.P. Gregory and R.A. Flavell

Butterworths: 1983. Pp.474. £52, \$99.95.

The explosion in our understanding of the structure and function of eukaryotic genes is nearly matched by the explosion in books describing it. However, it would seem that Maclean, Gregory and Flavell have produced one of the best accounts yet of the subject.

Far from being a disparate collection of short articles in incompatible styles (as is so often the case), the different reviews contained within the book have been carefully stitched together, in relatively homogeneous and concise format. Another valuable feature is that most of the material is up to date, to 1983, though such praise obviously applies more to some chapters than others. Before each section (group of related reviews), an introduction or summary has been provided. Some of these attempt to be enthusiastic to the extent of light-heartedness; for example: a "collection of DNA segments... has been called either a library (for the intellectuals) or a bank (for the businessman)". I idly wonder which editors refer to gene libraries or gene banks!

My detailed estimation of different sections in the book is as follows. In Section I on gene regulation (which would have been better named "Structural Aspects of DNA in Chromatin"), the first review by Jean Thomas is, to my mind, excellent; it told me all I felt I needed to know about this somewhat biophysical end of the subject. The other two reviews here, dealing with non-histone proteins and DNA modification, are somewhat less useful, though this probably reflects the fact that we know rather little about the involvement of, for example, HMG proteins and DNA methylation in gene regulation.

Section II might again have been more appropriately entitled, this time "Organization of 'Unusual' Chromosomes in Eukaryotes" rather than the general title "Organization of Chromosomes". However, the reviews on mitochondrial, lampbrush and polytene chromosomes are very helpful — so many molecular biologists are familiar with various "classic" experiments carried out using these systems but don't really understand their biology.

The third section covers experimental approaches. I found it somewhat disappointing. Why no specific articles on *in vitro* transcription, transient expression and stable transformation of cloned genes? Indeed, the only chapter I felt enthusiastic about here was that discussing the oocyte-egg system. Not that a review of teratocarcinoma and its use in studying gene expression isn't welcome; it is just that more generally-used systems should have had priority.

Section IV on specific gene systems is clearly excellent; while the order of the reviews is again a bit strange, there is here a wealth of information hard to glean in total from other sources. The book ends with a pathetically (and regrettably) short section on plant genes, but its brevity is hardly the fault of the editors. Taking away yeast, which in any case is barely a plant or indeed a eukaryote, very little progress has been made. Perhaps this will encourage more of us to turn to plants. After all, as Goldberg says in his review, plant genes have all the familiar boxes in

TATA, Cap, GT, AG, AATAAA.

so that one could still feel a "mainline molecular biologist".

Both advanced students and more experienced researchers will benefit from reading this book. I suspect it will find a sizeable audience, large enough for the publishers seriously to consider bringing out a cheaper edition. □

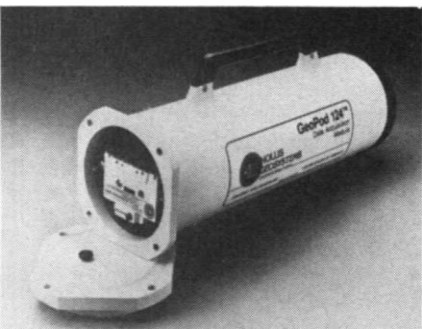
Nick Proudfoot is a University Lecturer and Tutor in Biochemistry at Brasenose College, University of Oxford.

From ecosystems to electrophoresis

New products covered this month include a completely sealed ecosystem being offered as a "gift item" and a computerized densitometer for electrophoresis.

● The use of linear roller bearing action in the new Shandon Hypercut rotary microtome is claimed to ensure a light, smooth cutting operation without sacrificing durability. This feature enables the operator to maintain a controlled cutting speed, producing sections of wax-embedded tissue blocks of an even, consistent thickness in a range from 0.5 to 30 μm . The Hypercut comes complete with a universal knife holder which features a unique lateral adjustment and accepts a wide range of knives, including glass, tungsten carbide tipped and disposable types.

Circle No. 100 on Reader Service Card.



For remote recording, the Hollis GeoPod.

● The GeoPod 124 data acquisition module is a long-term, magnetic tape cassette recording system designed for service in the field and in harsh industrial environments. Used in conjunction with its accompanying tape reader, the M-800, the GeoPod 124 is compatible with the IBM PC, DEC and other computing systems. Suggested uses for the GeoPod 124, from Hollis Geosystems, include meteorology, monitoring conditions on gas pipe-lines, hydrology, and structural strain measurements.

Circle No. 101 on Reader Service Card.

● New from Link Systems Ltd, the AN10000 microanalysis system is capable of measuring, storing, analysing, displaying and manipulating X-ray and electron signals produced by scanning and transmission electron microscopes. The basic 256K central processor is fast, flexible, and operates within a purpose-designed disk operating system which incorporates Winchester technology hard disks as a standard feature. AN10000 is equipped with a high resolution colour monitor, used in conjunction with a keyboard and MOUSE. The latter gives the operator a high degree of control and selection over displayed information.

Circle No. 102 on Reader Service Card.

● Simultaneous determination of nitrogen, carbon and sulphur is possible using two new versions of the Carlo Erba NA 1500 automatic nitrogen analyser. Samples of up to 100 mg (depending on sample composition) can be analysed. Other features are the short analysis times (3 min for N, 5 min for N and C, 6 min for S and 8 min for N, C and S), the wide measuring range (from 100 p.p.m. up to 100% for N, C and S) and high accuracy (better than 0.2%).

Circle No. 103 on Reader Service Card.

● Quickly changed, non-breakable, chemically inert fluoroplastic capillaries are used as the disposable tips for the new Oxford pipettes. Called the Ultra Micro, this series of five pipettes is designed to provide relatively high reproducibility and accuracy at 1, 2, 3, 4 or 5 μl . Both top and bottom of the internal air displacement chamber are sealed by O rings, for leak-free operation. Fluoroplastic (FEP) capillary tips can be used for tasks such as stirring a sample into a cuvette without fear of tip breakage. And it only takes a few seconds to change tips.

Circle No. 104 on Reader Service Card.

● The latest scanning electron microscope from the Electron Optics Group at Philips guarantees 5 nm resolution at 30 KV with 4 nm easily attainable. The SEM 515 attains higher resolution results from improved gun geometry, which gives a better signal-to-noise ratio. In a special low kv configuration this provides a fivefold increase in brightness when compared with the standard configuration. A new feature is the opportunity to include microprocessor-controlled automatic focusing and astigmatism correction. Based on algorithms developed by the Philips research laboratories in Hamburg, fully automatic focusing and astigmatism correction can be achieved in the most difficult circumstances.

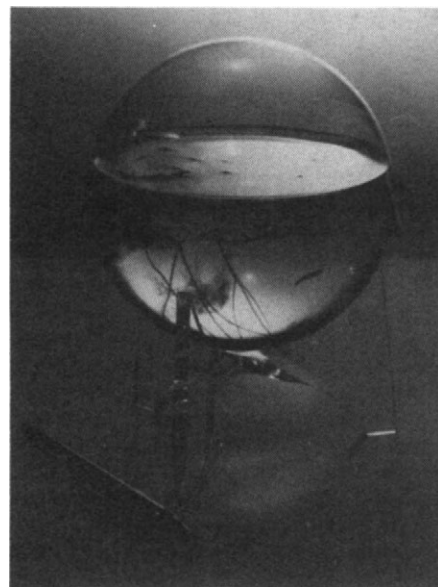
Circle No. 105 on Reader Service Card.

● Malvern Instruments' new Zetasizer II is designed to measure the surface charge on particles suspended in liquid media and also to analyse their size distribution. The sizer makes use of high-speed digital correlation analysis of the light scattered by the particles as they move under the influence of an electric field. The charge on the particles is automatically computed and displayed. Particle size measurement — in the range 10 nm to 3 μm — is based on analysis of Brownian motion of the particles.

Circle No. 106 on Reader Service Card.

● A digital storage oscilloscope, data acquisition and waveform analysis system — the Computerscope — is being offered by Advanced Instrumentation & Measurement Systems Ltd (AIMS) of Oxford. The system upgrades a general purpose microcomputer for use in high-speed data acquisition and analysis. The Computerscope includes all hardware and software necessary to convert a common microcomputer into a multi-purpose instrument at a much lower cost than traditional "stand alone" instruments that incorporate their own pre-programmed microprocessor. There are two models. The HR14 has up to four channels with a maximum sampling rate of 500 kHz at 14-bit resolution (0.006%) and Model D2 has two channels and a maximum sampling rate of 3.5 MHz at 8-bit resolution (0.4%). The Computerscope can also operate as a spectrum analyser or signal averager.

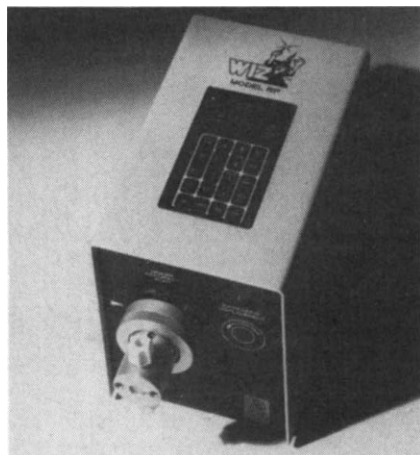
Circle No. 107 on Reader Service Card.



Shrimps and foliage share an EcoSphere.

● The glass globe above, little bigger than a softball, is an "EcoSphere". Engineering & Research Associates, Inc. (SEBRA) has taken a concept developed by space scientists and produced a "small-scale model of life on Earth". EcoSpheres are primarily intended as gifts but schools and museums have expressed interest in the EcoSphere as a teaching tool. The sphere contains a system including live shrimps, but the glass sphere is completely sealed, with no "breather" tube to supply exogenous oxygen.

Circle No. 108 on Reader Service Card.



Isco's new pump, the WIZ.

● WIZ RP from Isco is a chemically inert, high precision reciprocating pump which features microprocessor control. It has been designed for applications requiring precise metering at moderate pressures (100 p.s.i.) or for corrosive liquids and organic solvents which can damage peristaltic tubing. After initial calibration, any flow rate keyed in ml per hour (up to 7 litres per hour) will be produced with 1% accuracy. In the dispense mode, WIZ RP delivers any keyed-in volume. As a diluter, it takes up any volume of sample on a first push, and dispenses it followed by any volume of diluent on the second push.

Circle No. 109 on Reader Service Card.

ADVERTISEMENT

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem

☐ Other (please specify) _____

☐ I am interested in sending my order by mail.

☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

UMI Article Clearinghouse

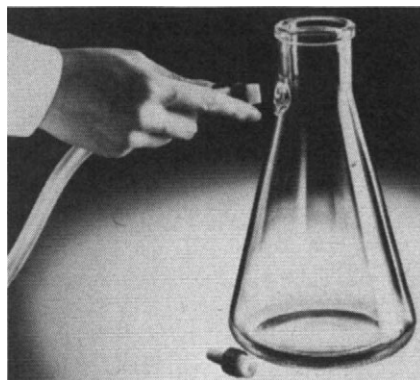
Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

● A new range of vacuum ovens, with volumes from 11 to 138 litres, is now available from Heraeus Equipment Ltd. Ovens in the new VTR series have an easy-to-clean cylindrical inner chamber which is built into a rectangular outer unit, allowing the ovens to be stacked. The large circular window in the door gives a clear view of the inner working chamber, the window itself being made of 15-mm toughened glass. The ovens feature an easy-to-read shelf thermometer, and shelves which can be heated individually are available as an option. These new ovens are designed specifically for sensitive materials which need to be dried at lower temperatures, offer shorter drying times and avoid decomposition or oxidation of the material being dried.

Circle No. 110 on Reader Service Card.

● A completely modular line of preparative scale columns and accessories intended to simplify scale-up of chromatographic separations has been introduced by LKB. The LKB 2237 preparative-scale columns and accessories enable the user to custom-design a preparatory or production-level system to handle working volumes of aqueous or organic solutions in the 0.2–10 litre range. Two series of 2237 columns are available — one for aqueous solvent and one for organic solvent. They enable all of the buffers and solvents required for even the most complex purification schemes, such as those for certain peptides, hormones and lipids, to be used. Columns are available in two diameters (5 and 10 cm) and three lengths (60, 100 and 120 cm) and are autoclavable.

Circle No. 111 on Reader Service Card.

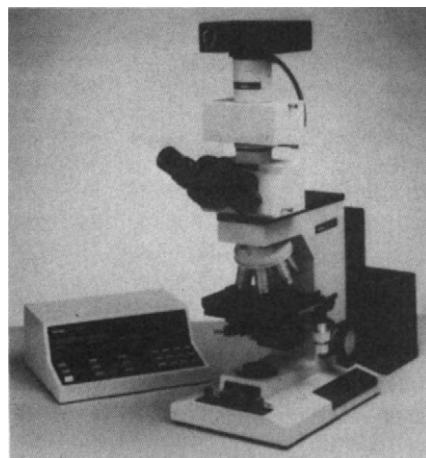


Detachable connection to a Bibby filter flask.

● Filter flasks in the new range from J. Bibby Science Products Ltd can be fitted with detachable connectors to enable tubing to be connected to the flasks without excessive hand pressure. The detachable tubing connector has a silicone rubber sleeve which pushes into the flask aperture. A screwcap then tightens up against the sleeve expanding it so that it provides a firm, leak-free seal. The new flasks are available in Pyrex borosilicate for general use or Quickfit for use with other Quickfit jointed ware. Capacities are from 100ml to 5 litres in Pyrex and 100ml to 1 litre in Quickfit.

Circle No. 112 on Reader Service Card.

© 1984 Nature Publishing Group



The Diastar pathologists' microscope.

● Diastar is a new pathologists' microscope from Reichert-Jung which can be used in conjunction with a new fully automatic, microprocessor controlled camera system, the Photostar. All functional and numerical data are entered via a keyboard into the microprocessor and alphanumeric LED displays are used. Photostar offers automatic film advance, multiple exposures on one frame; automatic matching of exposure to film characteristics; uniform background density on panoramic photographs and automatic exposure adjustment to take account of film or specimen characteristics. The Diastar can be adapted for multiviewing for up to five persons.

Circle No. 113 on Reader Service Card.

● Appraise is a computerized densitometer system for use with electrophoresis. Automated scanning provides programmable automatic step-over from track to track of all common electrophoresis media, and focused slit ultraviolet eliminates cross-talk and sample degradation. Appraise can also check calibration of absorbance and fluorescence. A detailed display of each pattern can be shown on the CRT for preview, and the unedited scan can be restored and displayed. Appraise, from Beckman, provides instant access to 10 programmable test selections and random access and viewing of 20 patterns without rescanning.

Circle No. 114 on Reader Service Card.

● Dynatech Laboratories' range of immunoassay plates includes conventional 96-well MicroElisa format plates ('U' or flat bottom) and also Removawell/Removastrip holders for use with either Removawells or Removastrips. These enable the operator to 'tailor-make' a plate using only as many wells or strips as necessary, and enable savings to be made in the use of antibody/antibody sensitization solutions.

Circle No. 115 on Reader Service Card.

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader service Card bound inside the journal.